

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

DIPLOMOVÁ PRÁCE

Brno, 2017

Bc. Miroslav Hynšt



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY

A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

**APLIKACE PRO VÝPOČET ŘEČOVÝCH PŘÍZNAKŮ
POPISUJÍCÍ HYPOKINETICKOU DYSARTRII**

APPLICATION FOR THE CALCULATION OF SPEECH FEATURES DESCRIBING HYPOKINETIC
DYSARTHRIA

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Miroslav Hynšt

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Tomáš Kiska

BRNO 2017

Diplomová práce

magisterský navazující studijní obor **Telekomunikační a informační technika**

Ústav telekomunikací

Student: Bc. Miroslav Hynšt

ID: 110540

Ročník: 2

Akademický rok: 2016/17

NÁZEV TÉMATU:

Aplikace pro výpočet řečových příznaků popisující hypokinetickou dysartrii

POKYNY PRO VYPRACOVÁNÍ:

V rámci této práce bude rozebrána problematika analýzy hypokinetické dysartrie, zejména u pacientů s Parkinsonovou nemocí. Dále bude navržena a implementována aplikace pro výpočet řečových příznaků, které popisují hypokinetickou dysartrii v oblasti fonace, artikulace plynulosti řeči či prozodie.

DOPORUČENÁ LITERATURA:

[1] MEKYSKA, J.; REKTOROVA, I.; SMEKAL, Z. Selection of Optimal Parameters for Automatic Analysis of Speech Disorders in Parkinson's Disease [online]. Telecommunications and Signal Processing (TSP), 2011 34th International Conference on, 2011: s. 408–412.

[2] RUSZ, J.; CMEJLA, R.; RUZICKOVA, H.; RUZICKA, E.: Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. J Acoust Soc Am, roč. 129, č. 1, 2011: s. 350–367.

Termín zadání: 1.2.2017

Termín odevzdání: 24.5.2017

Vedoucí práce: Ing. Tomáš Kiska

Konzultant:

doc. Ing. Jiří Mišurec, CSc.
předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

ABSTRAKT

Tato práce se zabývá návrhem a zhotovením aplikace pro výpočet řečových parametrů u lidí trpící Parkinsonovou nemocí. Na začátku je obecně popsána Parkinsonova nemoc a Hypokinetická dysartrie a jak ovlivňuje řeč a parametry řeči při jejím postižení. Rozebrány jsou hlavně oblasti jako fonace, prozodie, artikulace a plynulost řeči. Dále práce popisuje parametry s velkým významem při diagnostice této nemoci a jejím postupu. Zmíněny a popsány jsou i některé významné studie ve světě popisující průběh vyšetření řeči účastníků a výpočet některých řečových parametrů za účelem analýzy řeči lidí postižených Parkinsonovou nemocí. Součástí práce je navržená a implementovaná aplikace pro výpočet, export a vizualizaci řečových parametrů z vybraných zdrojových zvukových nahrávek.

KLÍČOVÁ SLOVA

Parkinsonova nemoc, Hypokinetická dysartrie, fonace, prozodie, artikulace, řečové parametry, Praat, Java

ABSTRACT

This thesis is about design and implementation of application for computing speech parameters on people with Parkinson disease. At the beginning is generally described Parkinson disease and Hypokinetic dysarthria and how it affects the speech and speech parameters when it occurs. Mainly there are described areas of speech like phonation, prosody, articulation and fluent speech. As a part of next topic this thesis describes specific speech parameters with bigger meaning during diagnosis Parkinson disease and it's progress over the time. There are also mentioned few significant studies dealing with examination of speech of the subjects with diagnoses of Parkinson disease and computing some speech parameters in order to analyze their speech impairments. Part of the thesis is description of implemented standalone application for calculating, exporting and visualizing of speech parameters from selected sound records.

KEYWORDS

Parkinson's disease, Hypokinetic dysarthria, phonation, prosody, articulation, speech parameters, Praat, Java

HYNŠT, Miroslav *Aplikace pro výpočet řečových příznaků popisující hypokinetickou dysarthrii*: diplomová práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2016. 56 s. Vedoucí práce byl Ing. Tomáš Kiska

PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma „Aplikace pro výpočet řečových příznaků popisující hypokinetickou dysartrii“ jsem vypracoval(a) samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor(ka) uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil(a) autorská práva třetích osob, zejména jsem nezasáhl(a) nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom(a) následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno

.....
podpis autora(-ky)

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu diplomové práce panu Ing. Tomáši Kiskovi, za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci.

Brno

.....

podpis autora(-ky)

OBSAH

Úvod	11
1 Parkinsnova nemoc a Hypokinetická dysartrie	12
1.1 Parkinsonova nemoc	12
1.2 Hypokinetická dysartrie	13
1.2.1 Fonace	13
1.2.2 Artikulace	13
1.2.3 Prozodie	14
1.2.4 Plynulost řeči	14
2 Parametry řeči	15
2.1 Frekvence základního tónu	15
2.2 Doba nástupu hlasu	16
2.3 Tempo řeči	16
2.4 Intenzita řečového signálu	17
2.5 DDK rychlost a pravidelnost	18
2.6 Formanty F1 a F2	18
2.7 Ukončení a navození fonace	19
3 Výsledky studií zaměřených na analýzu řeči u lidí s Parkinsonovou nemocí	20
3.1 Zkoumání doby nástupu hlasu	20
3.1.1 Cíl této studie	20
3.1.2 Účastníci	20
3.1.3 Průběh	21
3.1.4 Výsledky	22
3.2 Průběh dysprozodie u parkinsonovy nemoci	22
3.2.1 Cíl	22
3.2.2 Účastníci	23
3.2.3 Průběh	23
3.2.4 Výsledky	23
3.3 Akustická analýza obyčejné a konverzační řeči u jedinců s Parkinsonovou nemocí	24
3.3.1 Cíl	24
3.3.2 Účastníci	24
3.3.3 Průběh	24
3.3.4 Výsledky	25

4	Návrh systému analýzy řeči u pacientů s parkinsonovou nemocí	26
4.1	Popis konceptu systému	26
4.1.1	Test 3F: Dysartrický profil	26
4.1.2	Databáze nahrávek	27
4.1.3	Předzpracování	28
4.1.4	Extrakce parametrů	28
4.1.5	Export výsledků	29
4.1.6	Vizualizace	29
5	Popis systému	30
5.1	Použité technologie	30
5.1.1	Aplikace	30
5.1.2	Výpočet řečových parametrů	30
5.1.3	Vizualizace	31
5.1.4	Použité knihovny	31
5.1.5	Nápověda programu	31
5.2	Postup výpočtu parametrů	32
5.3	Hlavní komponenty programu grafického uživatelského rozhraní . . .	33
5.3.1	Nástrojová lišta	33
5.3.2	Seznam zvukových souborů	35
5.3.3	Hlavní panel s výsledky	36
5.3.4	Okno s hodnotami výpočtů	36
5.3.5	Visualizace výsledků	37
5.3.6	Stavová lišta	38
5.4	Hlavní funkce programu	38
5.4.1	Načtení zvukových souborů	38
5.4.2	Výpočet formantů	38
5.4.3	Výpočet F_0	40
5.4.4	Výpočet intensity	42
5.4.5	Spouštění externích skriptů	42
5.4.6	Nastavení programu	43
5.4.7	Nápověda	44
5.4.8	Logování programu	44
6	Souhrn vypočítaných parametrů	46
7	Závěr	50
	Literatura	51

Seznam symbolů, veličin a zkratk	54
Seznam příloh	55
A Obsah přiloženého CD	56

SEZNAM OBRÁZKŮ

4.1	Schéma systému pro analýzu nahrávek	26
4.2	Ukázka grafu prvních dvou formantů	29
5.1	Postup výpočtu	32
5.2	Hlavní okno aplikace	33
5.3	Nástrojová lišta	33
5.4	Seznam zvukových souborů	35
5.5	Vyhledávací pole	35
5.6	Tlačítka pro editaci seznamu nahrávek	36
5.7	Okno s výsledky	36
5.8	Grafické porovnání 2 výsledků	37
5.9	Stavová lišta	38
5.10	Detailní výsledky výpočtu formantů	39
5.11	Graf hodnot formantů	39
5.12	Graf pozic samohlásek	40
5.13	Detailní výsledky funkce <i>Voice report</i>	41
5.14	Graf hodnot F0	41
5.15	Detailní výsledky hodnot intenzity	42
5.16	Graf hodnot intenzity	43
5.17	Nastavení programu	44
5.18	Nápověda programu	45

SEZNAM TABULEK

3.1	Popis účastníků studie s parkinsonovou nemocí	21
3.2	Měřené parametry použité v této studii	23
4.1	Základní statistické údaje o databázi PARCZ	28
5.1	Výčet knihoven	32
5.2	Tlačítka v nástrojové liště a jejich ikony	34
6.1	Parametry popisující fonaci	47
6.2	Parametry popisující pauzování	47
6.3	Parametry popisující hybnost jazyka	49
6.4	Parametry popisující kvalitu řeči	49

ÚVOD

Popis řeči je velmi komplexní proces na který má vliv mnoho různých parametrů, které popisují mluvený projev. Některé mohou být ovlivněny věkem nebo pohlavím a nebo mnoha dalšími různými aspekty. Jedním z aspektů, které ovlivňují řeč je Parkinsonova nemoc a její důsledky na svaly řečového ústrojí. Jednou z poruch, která se v důsledku výskytu Parkinsonovy nemoci může vyskytnout je Hypokinetická dysartrie v jejímž důsledku postižení řeči mohou být také ovlivněny některé z těchto parametrů řeči a podle tohoto můžeme podrobnější analýzou například analyzovat průběh Parkinsonovy nemoci.

Tato práce se zabývá výpočtem řečových příznaků, které popisují Hypokinetickou dysartrii. Tyto příznaky mohou být využity k další analýze jako například různé statistické metody pro vyhodnocování stupně Parkinsonovy nemoci.

První kapitola stručně popisuje Parkinsonovu nemoc a její symptomy a dále Hypokinetickou dysartrii jako následek oslabení řečových svalů. Konkrétně jsou popsány oblasti řeči fonace, artikulace plynulosti řeči či prozodie a jak konkrétně je řeč ovlivněna v těchto oblastech.

V druhé kapitole jsou podrobněji popsány některé často používané parametry pro analýzu řeči. Výpočet mnoha z těchto parametrů je implementováno v praktické části této práce.

V další kapitole najdeme souhrn několika studií, které se zaměřily na výpočet parametrů řeči u lidí s parkinsonovou nemocí a jejich analýzu a jejich spojitost s PN. V každém popisu studie lze nalézt její cíle, popis účastníků, průběh a její výsledky.

Jak již bylo zmíněno cílem této práce je i aplikace pro výpočet některých parametrů popisujících řeč. V následující kapitole nalezneme její podrobnější návrh a detaily této aplikace. Součástí návrhu je schéma a rozbor jednotlivých hlavních funkcí tohoto systému. V rámci návrhu je popsána databáze PARCZ, která je použita jako zdroj nahrávek určených k analýze.

Výstupem této práce je implementovaná aplikace navržená také v rámci této práce a schopná vypočítat některé řečové parametry, zobrazit vypočítané hodnoty aji jako výčet hodnot a i graficky a získané výsledky případně exportovat. Program je možné rozšiřovat o mnoho dalších funkcí, z nichž některé jsou popsány v závěru práce.

1 PARKINSNOVA NEMOC A HYPOKINETICKÁ DYSARTRIE

Tato kapitola se zabývá teoretickým rozbořem Parkinsonovy nemoci a některých jejích příznaků. Zejména pak oblasti ovlivňujících řeč tzv. Hypokinetickou dysartrií. Dále jsou zde popsány některé jednotlivé oblasti řeči, u kterých se hypokinetická dysartrie projevuje.

1.1 Parkinsonova nemoc

Parkinsonova nemoc byla poprvé popsána anglickým lékařem Jamesem Parkinsonem v roce 1817 v jeho práci „An Essay on the Shaking Palsy“. Přesná příčina výskytu do dnes není známa i když stálými faktory výskytu parkinsonovy nemoci jsou stáří a genetika (výskyt v rodině). Je to jedna z nejčastějších neurodegenerativních poruch, která je charakterizována rychlým odumíráním dopaminergních nervových buněk v oblasti středního mozku. [3, 5, 6]

Přibližně 70% pacientů postiženými parkinsonovou nemocí trpí poruchami řeči. PN se projevuje mnoha různými příznaky ať už motorickými či nemotorickými.

Mezi základní **motorické příznaky** PN, které ovlivňují řeč do různé míry, patří: [4]

- **Bradykineze** - zpomalenost, redukováná rychlost pohybu
- **Rigidita** - ztuhlost svalů
- **Tremor** - třes
- **Akineze** - snížený start pohybů

Mezi další **nemotorické příznaky** patří:

- **Spánkové poruchy** - nespavost, poruchy REM spánku, rozsáhlá denní ospalost
- **Neuropsychiatrické rysy** - apatie, úzkost, deprese případně i halucinace
- **Smyslová dysfunkce** - ztráta čichu, snížení vnímání kontrastu a rozlišování barev
- **Bolest**
- **Únava**

1.2 Hypokinetická dysartrie

V předešlé části jsme obecně rozebrali parkinsonovu nemoc a její běžné příznaky. Nyní se zaměříme na konkrétní příznaky související s poruchami řeči a to konkrétně některé oblasti Hypokinetické dysartrie.

Obecně pojmem dysartrie rozumíme poruchu řeči. Hypokinetickou dysartrií (HD) označujeme motorické poruchy řeči kdy dochází k oslabení svalové kontroly řečových mechanismů. I když se HD vyskytuje u lidí s parkinsonovou nemocí velmi často, tak ne vždy je toto pravidlem. Projevy HD se tradičně spojují s neurologickými dopaminergně vázanými symptomy PN jako jsou rigidita, hypokineze, bradykineze, akineze a klidový tremor. [2]

Konkrétní symptomy běžně spojené s hypokinetickou dysartrií zahrnují snížený důraz při výslovnosti, nepřesné souhlásky, nevhodné pauzy, špatná kvalita nezvučného hlasu. [6]

Z hlediska zpracování řečového signálu a narušené oblasti řeči v rámci HD se v této práci zaměříme hlavně na tyto dílčí složky řeči: *fonace*, *artikulace*, *prozodie* nebo *plynulost řeči*.

1.2.1 Fonace

Fonace bývá zpravidla první oblastí kde se Hypokinetická dysartrie projevuje. [4] Nejvíce charakteristické vlastnosti poruch řeči spojeny s parkinsonovou nemocí jsou spojeny s fonací. [14] Obecně poruchy v této oblasti jsou způsobeny ztuhlostí svalů v oblasti hrtanu a tedy i ztuhlostí hlasivek. Jako důsledek této ztuhlosti u pacientů registrujeme významné zvýšení frekvence základního tónu F_0 při vyslovování dlouhých samohlásek.

Dalším symptomem v oblasti fonace je doba nástupu hlasu, která je definována jako doba mezi zazněním vyslovené souhlásky a hlasového nástupu následující samohlásky. Tato doba se liší u pacientů postižených Parkinsonovou nemocí v porovnání se zdravými lidmi. Může být buďto kratší nebo delší.

Rigidita také vede k narušení procesu respirace, který slouží k produkování toku a tlaku vzduchu důležitého pro řeč. Toto může také ovlivnit řečnickovu schopnost mluvit normální hlasitostí a ovládat ji. [6]

Rozsah intenzity se také snižuje u jednotlivců s PN a toto se připisuje mimo jiné také narušení procesu respirace.

1.2.2 Artikulace

Schopnost artikulace úzce souvisí s činností svalů v oblasti úst a obličeje. Fyziologické studie artikulačního svalstva ukazují na snížení rozsahu a rychlosti pohybu těchto

svalů při postižení PN. Zejména je narušena výslovnost souhlásek. [2]

Mnoho výzkumníků reportovalo, že tzv. plovivní souhlásky neboli *ploviva* (např. p, t, k, b, d, g) byly u pacientů s Parkinsonovou nemocí vyslovovány nepřesně a spíše zněly jako *frikativa* (např. f, v). Normální mluvčí ukazovali výrazné snížení síly v momentě zavírání úst, kdežto pacienti s PN neukázali žádné snížení síly při zavírání úst a tedy špatné výslovnosti.

Jednou z metod vyhodnocování artikulačních schopností u pacientů s PN je využití *diadochokinetických úloh* DDK (diadochokinetic tasks). Tyto úlohy typicky zahrnují vyslovování řetězců slabik obsahující kombinaci souhlásky a samohlásky např. *pa-pa-pa*, *ta-ta-ta* nebo *ka-ka-ka* a nebo i jejich kombinace jako *pa-ta-ka* apod. Vyhodnocení DDK úloh se používají k vyhodnocování pacientovi schopnosti provádět rapidně se měnící artikulační pohyby.

Produkce samohlásek může být také ovlivněna kvůli špatnému rozsahu a rychlosti pohybu rtů a čelisti. [6]

1.2.3 Prozodie

Prozodie jako rytmus a melodie řeči je také velmi důležitá jako příznak PN. Mezi tři základní prozodické rysy patří kmitočet základního tónu, intenzita a tempo řeči. Kromě těchto základních parametrů mezi prozodické rysy také patří např. rytmus, intonace, přízvuk aj.

Frekvence základního tónu se u lidí s PN liší bez ohledu na to zda-li je věta oznamovací nebo tázací. Co se intenzity týče, tak pacienti s PN demonstrují sníženou schopnost měnit intenzitu hlasu ke zvýraznění semantiky věty nebo emocí.

Tempo nebo rychlost řeči může být u pacientů s PN zrychlené (*bradyfemie*) nebo zpomalené (*tachyfemie*). Zrychlené tempo může být způsobeno artikulačními problémy s nepřesnou výslovností a tedy způsobující zrychlení řeči. Tachyfemie zhoršuje artikulaci, má vliv na umísťování pauz v mluveném projevu a významně ovlivňuje celkovou srozumitelnost řeči. [2] Vzhledem k tomu že tato rychlost se může pohybovat oběma směry, význam těchto rozdílů není tak důležitý. [6]

1.2.4 Plynulost řeči

Z hlediska plynulosti se u lidí s PN objevují dvě formy poruch, které se mohou vyskytnout současně nebo pouze izolovaně a to jsou *palilalie* a *hezitace*. *Palilalie* označuje zrychlené opakování slov či jednotlivých slabik na začátku, uprostřed nebo konci slova. Co se *hezitace* týče, označuje tato porucha neúmyslné pauzy při zahajování řeči a nebo kdykoliv během mluvení. [4]

2 PARAMETRY ŘEČI

V této kapitole se budeme zabývat různými parametry popisující řečové poruchy zmíněné výše v sekci o Hypokinetické dysartrii. Tyto parametry mohou být dále využity k automatickému zpracovávání řečových záznamů. Mohou se použít v analýze Hypokinetické dysartrie a pro další vyhodnocování jako například statistická analýza řeči.

2.1 Frekvence základního tónu

Jedním z nejzákladnějších parametrů pomocí kterých můžeme zkoumat poruchy řeči u lidí s Parkinsonovou nemocí je frekvence základního kmitočtu F_0 . U těchto lidí můžeme pozorovat zvýšení této frekvence jak pro delší mluvené části (např. slova) tak pro dlouhé samohlásky. Je známo, že s delším časovým postižením Parkinsonovou nemocí se střední hodnota $mean(F_0)$ této frekvence zvyšuje, a proto může být použita k určení vážnosti nemoci. [7]

Standartně se využívá měření střední hodnoty $mean(F_0)$, směrodatné odchylky F_0 SD a variační rozsah F_0 VR. I přesto, že se našli významné rozdíly u absolutních hodnot F_0 a jejího rozsahu mezi pacienty s PN a zdravými lidmi, jsou tyto hodnoty ovlivněny individuálními rozdíly mezi lidmi jako například pohlaví. Specificky rozsah F_0 se vztahuje k průměrné intenzitě hlasu daného jedince. [8] Z tohoto důvodu jsou lepším ukazatelem relativní parametry z nichž některé jsou popsány níže.

Dalším užitečným parametrem může být *variační rozsah* F_0 VR (F_0 Variation range) definovaný jako:

$$F_0 \text{ VR} = F_{0max} - F_{0min} \text{ (Hz)} \quad (2.1)$$

Dále se dá pomocí F_0 a $meanF_0$ vypočítat další parametry jako $relF_0$ SD a $relF_0$ VR. [9]

$relF_0$ SD značí vztah mezi relativní směrodatnou odchylkou frekvence základního tónu a $meanF_0$:

$$relF_0 \text{ SD} = \frac{F_0 \text{ SD}}{meanF_0} \cdot 100 (\%) \quad (2.2)$$

$relF_0$ VR značí vztah mezi relativní variační rozsah frekvence základního tónu a $meanF_0$:

$$relF_0 \text{ VR} = \frac{F_0 \text{ VR}}{meanF_0} \cdot 100 (\%) \quad (2.3)$$

Pro získání F_0 existuje spousta metod mezi, které patří například *autokorelační funkce* (ACF), *křížová korelace* (CCF) nebo *detektory základního tónu* PDA.

Jedním z parametrů vyjadřujícím variabilitu základní frekvence F_0 v čase je frekvenční nestabilita neboli *jitter*. Definuje se jako rozdíl v délce dvou sousedních period hlasivkových pulzů v porovnání s průměrnou délkou periody.

Existuje mnoho variant lišící se způsobem výpočtu. Jsou to například PPQ5 (five-point Pitch Perturbation Quotient), RAP (Relative Average Perturbation), jitt loc (průměrná absolutní odchylka mezi následujícími periodami dělená průměrnou periodou), jitt abs (průměrná absolutní odchylka mezi následujícími periodami), jitt ddp (průměrná absolutní odchylka mezi následujícími odchylkami sousedních glotálních period dělená průměrnou periodou). [17]

Parametr popisující kolísání úrovně hlasivkových pulzů se nazývá *shimmer*. Zatím co jitter popisuje změny periody, shimmer představuje změny amplitudy hlasivkových pulzů.

2.2 Doba nástupu hlasu

Doba nástupu hlasu VOT (Voice onset time) je definována jako doba od dokončení vyslovení plovivní souhlásky nebo-li *ploviva* do zvukového nástupu následující samohlásky.

Tato doba může být kladná, záporná nebo nulová a závisí také na druhu následující samohlásky. Například pro *bilabiální* souhlásky (b, p) je doba nástupu hlasu relativně krátká (kolem 25 ms), v případě *alveolárních* souhlásek (d, t) se jedná o středně dlouhou dobu a u *velárních* dosahuje tato doba nejvyšších hodnot (až 100 ms). [4]

Měření VOT by mohlo sloužit jako ukazatel Parkinsonovy nemoci ale mnoho publikací poukazuje na to, že tato doba není zcela jednoznačným ukazatelem. Přesné určení hranic VOT je celkem komplikované a neexistuje žádný spolehlivý algoritmus na přesné měření této doby. [11]

Další obdobou tohoto parametru může být *poměr doby nástupu hlasu* (VOT ratio), který je definovaný následující rovnicí: [13]

$$\text{VOT ratio} = \frac{\text{VOT}}{\text{trvání slova}} \quad (2.4)$$

2.3 Tempo řeči

Jak již bylo zmíněno v 1.2.3 tempo řeči je jedním ze základních prozodických rysů. K analyzování tempa se dá využít několika parametrů. Níže budeme definovat některé často používané v analýze řeči. Doba pauzy se dá definovat různě ale většinou se definuje jako doba ticha přesahující okolo 60 ms.

- **TST (total speech time)** - celková doba řeči (s)

- **TPT (total pause time)** - celková doba pauz v řeči (s)
- **NST (net speech time)** - $TST - TPT$, doba čistě řečového signálu (s)
- **TSR (total speech rate)** - *celkové tempo řeči* definováno jako poměr celkového počtu slabik a TST (slabiky/s)
- **NSR (net speech rate)** - *tempo čistě řečového signálu* definováno jako poměr celkového počtu slabik a NST ($TST - TPT$) (slabiky/s)
- **PR (pause ratio)** - *poměr pauz* definováno jako procentuální poměr celkové doby všech pauz a TST

$$\frac{TST - NST}{TST} \cdot 100 (\%) \quad (2.5)$$

Při různých měření bylo použito mnoho dalších parametrů z hlediska prozodie mezi které patří například: [10, 11]

- **PallN** - počet všech pauz
- **PinwT** - doba všech pauz uvnitř víceslabičných slov
- **PinwN** - počet všech pauz uvnitř víceslabičných slov
- **PewT** - doba všech pauz na konci slov
- **PewN** - počet všech pauz na konci slov
- **Pew (%)** - procento všech pauz na konci slov
- **PRinw (%)** - procento všech pauz uvnitř víceslabičných slov z celkového počtu pauz
- **VKT (vowel keeping time)** - doba znění samohlásky, co nejdelší vyslovení samohlásky na jeden dech

2.4 Intenzita řečového signálu

Intenzita řeči je dalším ukazatelem využívaným v analýze v oblasti prozodie. Ačkoliv je intenzita neboli hlasitost velmi specifický parametr, protože závisí na mnoha specifických a individuálních vlastnostech člověka, i přesto lze využít k analýze. Lidé mohou mluvit s různou intenzitou hlasu i na základě emocí například když jsou naštvaní nebo smutní a dočasně tak intenzitu svého hlasu zvýšit ale jen dočasně. Většina publikací se shoduje, že Parkinsonova nemoc snižuje u pacientů jejich intenzitu hlasu.

Ukázalo se, že pacienti s Parkinsonovou nemocí mají celkově nižší stupeň intenzity, nedostatky v rozsahu intenzity a nižší směrodatnou odchylku během řeči. Vzhledem k tomu, že intenzita se vztahuje k individuálním vlastnostem jednotlivců jako například pohlaví tak se nepoužívají absolutní hodnoty stupně intenzity nebo absolutní rozsah intenzity, ale spíše relativní měření. Jedna z podmínek úspěšného měření je zajištění stejné vzdálenosti od mikrofonu pro všechny nahrávky. [11]

Tuto intenzitu je možné z hlediska zpracování signálů vyjádřit pomocí krátkodobé energie E nebo prostřednictvím Teagrova energetického operátoru TEO (Teager Energy Operator). Fluktuace intenzity může být vyjádřena směrodatnou odchylkou nebo pomocí parametru *shimmer*, které měří stupeň změny signálu v čase. [4]

2.5 DDK rychlost a pravidelnost

DDK úlohy jsou zmíněny již v sekci o poruchách v artikulaci 1.2.2. Mají za úkol zjistit schopnost pacienta opakovat rychle a stálou rychlostí kombinace souhlásek a samohlásek. Průměr DDK rychlosti určuje počet promluvených slabik za sekundu. Koeficient DDK pravidelnosti měří stupeň proměnlivosti DDK hodnoty za jednotku času a vyhodnocuje schopnost udržovat stálou rychlost říkání kombinací souhlásek a samohlásek. [11]

2.6 Formanty F1 a F2

Při mluvení vznikají v ústních dutinách rezonancí zesílené frekvence, které nazýváme *formanty* (F_1 , F_2 , F_3).

Hodnoty těchto frekvencí jsou ovlivňovány pozicí řečových orgánů jako třeba jazyka. Díky tomu můžeme pomocí formantů sledovat například pohyblivost jazyka. První dva formanty F_1 a F_2 jsou parametry, které odhalí nejvíce informací o vytváření samohlásek a na základě charakteristik odvozených z tohoto faktu jsou velmi často použity k popisu řeči u pacientů s Parkinsonovou nemocí. Běžně se s formanty spojují i odpovídající šířky pásma B_1 a B_2 . Velká nevýhoda formantů je jejich závislost na věku a pohlaví. Jak se již zmínilo můžeme pomocí formantů měřit pohyb jazyka na čemž je založen parametr *VSA* (Vowel Space Area). Při výpočtu tohoto parametru jsou nejprve extrahovány samohlásky z řeči a poté jsou odhadnuty jejich první dva formanty F_1 a F_2 . Tyto hodnoty poté co jsou vykresleny tvoří pentaagon jehož obsah odpovídá parametru *VSA*. Výpočet tohoto parametru je popsán rovnicemi definovanými níže: [1]

$$|v_{xy}| = \sqrt{(F_{1x} - F_{1y})^2 + (F_{2x} - F_{2y})^2} \quad (2.6)$$

$$a_1 = |v_{ue}|, b_1 = |v_{ei}|, c_1 = |v_{iu}| \quad (2.7)$$

$$a_2 = |v_{uo}|, b_2 = |v_{oe}|, c_2 = |v_{eu}| \quad (2.8)$$

$$a_3 = |v_{oa}|, b_3 = |v_{ae}|, c_3 = |v_{eo}| \quad (2.9)$$

$$s_i = \frac{a_i + b_i + c_i}{2} \quad (2.10)$$

$$Tr_i = \sqrt{s_i(s_i - a_i)(s_i - b_i)(s_i - c_i)} \quad (2.11)$$

$$VSA = \sum_{i=1}^3 Tr_i (Hz^2) \quad (2.12)$$

F_{ix} a F_{iy} vyjadřují frekvence i -tého formantu odpovídající samohlásky.

Dále lze použít parametr lnVSA, který se počítá jako VSA, ale všechny hodnoty jsou počítány pomocí logaritmu. U pacientů postižených Parkinsonovou nemocí se očekává, že tento parametr bude nabývat menších hodnot, ale i tento parametr je opět velmi závislý na věku a pohlaví. Kvůli tomuto se musí zavést určité normalizace, které by potlačily tyto rozdíly. Jedním z možných navržených parametrů je VAI (Vowel Articulation index) a dalším takovým parametrem může být FCR (Formant Centralization ratio), který je vyjádřen jako:

$$FCR = \frac{F_{2u} + F_{2a} + F_{1i} + F_{1u}}{F_{2i} + F_{1a}} \quad (2.13)$$

Normalizace je zde provedena dělením součtem formantů F_{2i} a F_{1a} hlásek a a i . [1]

2.7 Ukončení a navození fonace

Bylo zjištěno, že pacienti s Parkinsonovou nemocí mají problémy s náhlým ukončením znělého úseku. Naopak začátek znělého úseku se příliš neliší od normálních lidí. Tento parametr popisuje změnu frekvence základního tónu F_0 . Část ukončení fonace se označuje posledními 10 hlasivkovými pulzy ve znělém úseku. Druhá část navození fonace se označuje také 10 hlasivkovými pulzy na začátku znělého úseku. Perioda mezi prvním a druhým pulzem je značena T_0^1 což je převrácená hodnota F_0^1 . F_0^2 může být vyjádřena stejným způsobem (je to perioda mezi druhým a třetím pulzem). Ke sledování změn mezi těmito frekvencemi se vyjadřují tyto frekvence pomocí půltónů (Semitone ST):

$$ST = 12 \cdot \log_2\left(\frac{F_0^x}{F_0^1}\right) \quad (2.14)$$

V analýze se potom sleduje dynamika změny těchto frekvencí. [4]

3 VÝSLEDKY STUDIÍ ZAMĚŘENÝCH NA ANALÝZU ŘEČI U LIDÍ S PARKINSONOVOU NEMOCÍ

V této kapitole je popsáno několik publikovaných studií, které vyšetřují různé parametry používané při analýze poruch řeči u pacientů s Parkinsonovou nemocí. U každé studie je popsán co je jejím cílem, její průběh, účastníci studie a její výsledky.

3.1 Zkoumání doby nástupu hlasu

Parametrem doby nástupu hlasu VOT bylo již rozebíráno v mnoha publikacích hlavně z důvodu prokázání jeho spojitosti parkinsonovou nemocí. Tato studie byla provedena Emily Fischer a Alexander M. Goberman v roce 2009. [13]

3.1.1 Cíl této studie

Cílem této studie bylo prozkoumat dobu nástupu hlasu v projevu jednotlivců s parkinsonovou nemocí založené na dvou různých měření:

- běžné měření doby nástupu hlasu VOT
- VOT bez efektu rychlosti tzv. VOT poměr

Tyto měření byly použity k zodpovězení následujících otázek:

- Existuje rozdíl mezi VOT a VOT poměrem u porovnání jednotlivců s PN a věkově odpovídajících účastníků bez PN?
- Má výška samohlásky a artikulační poloha souhlásky vliv na VOT nebo VOT poměr?
- Existují změny VOT nebo VOT poměru v různých stádiích léčby parkinsonovy nemoci?

3.1.2 Účastníci

Této studii se účastnilo 6 mužů a 4 ženy ve věku od 48 do 80 let. Všichni účastníci splňovali následující kritéria:

- diagnostikováni parkinsonovou nemocí
- léčeni přípravky NAKOM (tablety použité při léčení PN)
- hlášeny výkyvy v reakci na léčbu
- rodilý mluvčí americké angličtiny
- žádné nahlášené neurologické operace nebo další neurologické nemoci
- žádné předchozí poruchy řeči, jazyka nebo sluchu nesouvisející s PN symptomy

Jeden z účastníků byl vyřazen kvůli nemožnému měření daných parametrů z důvodů pokročilosti jeho dysartrie. Podrobnější popis jednotlivých účastníků studie je ukázán v tabulce 3.1.2, přičemž efektivita léčby byla hodnocena na stupnici od 1 do 7, kde 1 = bez efektu a 7 = perfektní účinek. Podobná stupnice byla použita při hodnocení závažnosti symptomů, kde 1 = extrémní a 7 = vůbec závažné.

Účastníci	Věk	Pohlaví	Léta po diagnóze	Efektivita léčby		Vážnost symptomů		Vážnost HD bez léčby
				Bez léčby	Počas léčby	Bez léčby	Počas léčby	
1	48	Ž	7	1	6	3	6	Mírné
2	51	Ž	4	4	4	2	4	Mírné
3	69	Ž	11	1	7	4	6	Mírné
4	80	Ž	10	7	5	6	6	Mírné
5	71	M	9	6	6	4	4	Mírné
6	63	M	18	2	6	3	4	Střední
7	69	M	6	6	7	6	6	Střední
8	76	M	7	6	7	5	6	Střední
9	64	M	15	3	4	1	5	Závažné
10 ¹	74	M	8	5	5	6	6	Závažné
Průměr	65.6		9.7	4.0	5.8	3.8	5.2	

Tab. 3.1: Popis účastníků studie s parkinsonovou nemocí

3.1.3 Průběh

Každý účastník trpící PN byl dvakrát nahráván u něj doma. První nahrávka byla nahrána 30 minut před jejich ranní léčbou. Tyto nahrávky reprezentovaly stav „bez léčby“ a účastníci poskytli jejich vlastní odhad efektivy léčby a závažnost symptomů.

Druhá nahrávka se odehrála ve stejný den, 60 minut po jejich ranní léčbě. Tyto nahrávky reprezentovaly stav „počas léčby“ a opět poskytli jejich vlastní odhady efektivy léčby a závažnosti symptomů.

Na začátku prvního nahrávání dostali účastníci 3 různé úkoly, které měli ohodnotit jejich závažnost a typ dysartrie. První úkol bylo čtení Rainbow passage² textu, který se používá k testování řečových schopností. Druhý úkol byl série DDK úloh

¹Jak bylo již zmíněno výše v textu tento účastník byl vyřazen.

²The Rainbow Passage text lze nalézt na straně 127 druhého vydání Voice and Articulation Drillbook

u nichž bylo úkolem mluvit co nejrychleji a pravidelně jak je to jen možné. Posledním úkolem bylo prodloužení výslovnosti samohlásky *a* tak dlouho a stejně tak jak je to jen možné na jeden nádech. Na základě těchto tří úkolů byl vyhodnocen typ a závažnost dysartrie certifikovaným *řečově-jazykovým patologem* SLP (Speech-Language Pathologist).

Pro nahrávání byl použit mikrofón Shure SM-58 a digitální audio rekordér Sony PCM-M1. Vzdálenost od mikrofónu při nahrávání byla 15 cm. Nahrávky byly analyzovány pomocí programu PRAAT.

3.1.4 Výsledky

Hodnocení typu a závažnosti dysartrie bylo provedeno z nahrávek ve stavu bez léčby aby se určilo jestli účastníci trpí Hypokinetickou dysartrií, která je typická pro pacienty s parkinsonovou nemocí. Výsledky ukázali, že všichni účastníci prokazují hypokinetickou dysartrií v jejich stavu bez léčby a že závažnost HD se pohybovalo od mírné do závažné (5 - mírná, 3 - střední a 1 - závažná).

Současná zjištění, že artikulační místo a výška samohlásky má vliv na dobu nástupu hlasu (VOT) souhlasí s předchozími studiemi. Tato studie však přidala fakt, že to platí pro pacienty s Parkinsonovou nemocí. Zjistila se, že výška samohlásky měla vliv na poměr doby nástupu hlasu (VOT ratio) ale ne na běžnou dobu nástupu hlasu (VOT). Toto potvrzuje, že změny doby nástupu hlasu jsou spojeny se změnami výšky samohlásky a že tento fakt nebyl výsledkem změn v rychlosti artikulace.

Celkově data z této studie podporují fakt, že zkoumání jak doby nástupu hlasu tak poměru doby nástupu hlasu je užitečné u lidí s PN, protože se ukázala disociace rychlostně závislých změn doby nástupu hlasu a změny pravé doby nástupu hlasu.

3.2 Průběh dysprozodie u parkinsonovy nemoci

Autoři této studie jsou Sabine Skodda, Heiko Rinsche a Uwe Schlegel. [9] I přesto, že dysprozodie je velmi dobře definována (změna tempa řeči a pauzování a další příznaky), jen málo se ví o těchto změnách v průběhu delšího časového úseku což si autoři vzali jako zaměření pro tuto studii.

3.2.1 Cíl

Cílem této studie je analyzovat změny rychlosti řeči a kolísání výšek u pacientů s Parkinsonovou nemocí v průběhu času a porovnání těchto výsledků se zdravými lidmi.

3.2.2 Účastníci

Této studii se účastnilo 50 pacientů s Parkinsonovou nemocí, kteří byli klinicky diagnostikováni, a 50 kontrolních zdravých lidí na srovnání. Věk pacientů se pohyboval od 43 do 80 let a z těchto 50ti pacientů bylo 27 mužů a 23 žen. Lidé z kontrolní skupiny byli věkově shodní a v této skupině bylo 25 žen a 25 mužů. Parkinsonova nemoc byla u pacientů diagnostikována od 1 do 20 let před prvním zkoumáním. Doba mezi prvním a druhým zkoumáním se pohybovala mezi 7 a 79 měsíci. Při obou zkoumáních prošli pacienti neurologickým vyšetřením podle UPDRS motorické stupnice. V době zkoumání užívali pacienti léky aspoň 4 týdny před zkoumáním.

3.2.3 Průběh

Každý účastník měl za úkol provést standardizovaný proslov skládající se ze čtení pasáže 4 komplexních vět. Po tomto prvním testu byl druhý pokus brán k hlubší analýze za účelem minimalizace problémů běžného čtení. Nahrávky byly digitálně nahrány za použití audio softwaru Steinberg WaveLab a mikrofону se sluchátkami Plantronics Audio 550 DSP.

V tabulce 3.2.3 jsou zobrazeny parametry naměřené během testování účastníků a použité v analýze jejich nahraných záznamů.

Zkratka	Definice
$\text{mean}F_0$ (Hz)	Průměrná frekvence základního tónu
F_0 SD (Hz)	Směrodatná odchylka frekvence základního tónu
F_0 VR (Hz)	Variační rozsah frekvence základního tónu
$\text{rel}F_0$ SD (%)	Relativní směrodatná odchylka F_0 vztažena k $\text{mean}F_0$
$\text{rel}F_0$ VR (%)	Variační rozsah relativní F_0 vztažen k $\text{mean}F_0$
TSR (slab./s)	Celková doba řeči
NSR (slab./s)	Tempo čistě řečového signálu
PR (%)	Procento pauz
Pinw (%)	Poměr pauz v rámci slova

Tab. 3.2: Měřené parametry použité v této studii

3.2.4 Výsledky

Jako výsledek této studie byla demonstrace výskytu vzoru rychlosti řeči u lidí s Parkinsonovou nemocí, artikulačním zrychlením v počátečních stádiích a zpomalením

během průběhu nemoci zvláště u mužských pacientů. Dále proměnlivost F_0 prokazovala zhoršení v průběhu času alespoň u žen. Jelikož doba trvání parkinsonovy nemoci u účastníků a doba mezi oběma testy nebyla nijak standardizována nebylo možné určit zda postup dysprozie následuje tempo zhoršování motoriky. Dodatečně ze studie vyplývá, že zhoršování řeči je individuální, protože se nenašla spojitost mezi změnami řečových parametrů a dobou mezi oběma testováními.

3.3 Akustická analýza obyčejné a konverzační řeči u jedinců s Parkinsonovou nemocí

Autory této studie jsou Alexandr M. Goberman a Lawrence W. Elmer. [15] Již mnoho studií bylo zaměřeno na porovnání obyčejné řeči versus konverzační řeči u lidí s řečovým postižením. Hlavním účelem předchozích studií bylo pomoci zvýšit srozumitelnost řeči kvůli lidem s problémy se sluchem.

3.3.1 Cíl

Cílem této studie bylo zjistit rozdíly mezi konverzační a obyčejnou řečí u jednotlivců s parkinsonovou nemocí.

3.3.2 Účastníci

Pro účely této studie posloužilo 12 jedinců (6 mužů a 6 žen). Jejich průměrný věk bylo 73 let (věkový rozsah 55 - 84 let) a všichni byli diagnostikováni jejich neurologem Parkinsonovou nemocí. Všichni užívali léky proti PN a jimi udávaná efektivita těchto léčiv byla v průměru 3.83 (rozsah 3 - 5) na pětibodové stupnici (1 = neúčinkuje, 3 = účinkuje středně, 5 = účinkuje velmi dobře). Žádný z účastníků neprošel řečovou nebo jazykovou terapií a žádný z nich netrpěl jiným neurologickým onemocněním, které by ovlivňovalo řeč. Všichni byli postiženi Hypokinetickou dysartrií různé závažnosti od mírné po závažnou.

3.3.3 Průběh

Prvním úkolem byla série $h + \text{samohláska} + d$ (hVd) vložené v anglické frázi „Say hVd again“. Druhým úkolem bylo čtení prvního odstavce Rainbow passage a třetím úkolem bylo provést dvouminutový monolog. Celá nahrávka trvala 15 minut pro jednoho účastníka. K analýze byly použity parametry jako $\text{mean}F_0$, F_0 SD, artikulační rychlost, poměr pauz, první a druhý formant a další.

3.3.4 Výsledky

Byl zaznamenán významný rozdíl v artikulační rychlosti při porovnání konverzační a obyčejné řeči ačkoliv nebyl zjištěn rozdíl pro artikulační rychlost u čtených versus mologových úloh.

Analýza jednotlivých účastníků objevila, že méně než polovina z nich vykazovala zvýšený poměr pauz v úlohách obyčejného monologu i přesto, že většina účastníků vykazovala zvýšený poměr pauz u úloh obyčejného čtení.

Testování $\text{mean}F_0$ objevilo významný rozdíl mezi konverzační a jasnou řečí, kdežto mezi mologovým čtením a čtením pasáže nebyl zjištěn žádný významný rozdíl. Analýza $\text{mean}F_0$ jednotlivých účastníků zjistila, že většina z nich vykazovala vyšší hodnoty pro obyčejnou řeč u čtení i monologu.

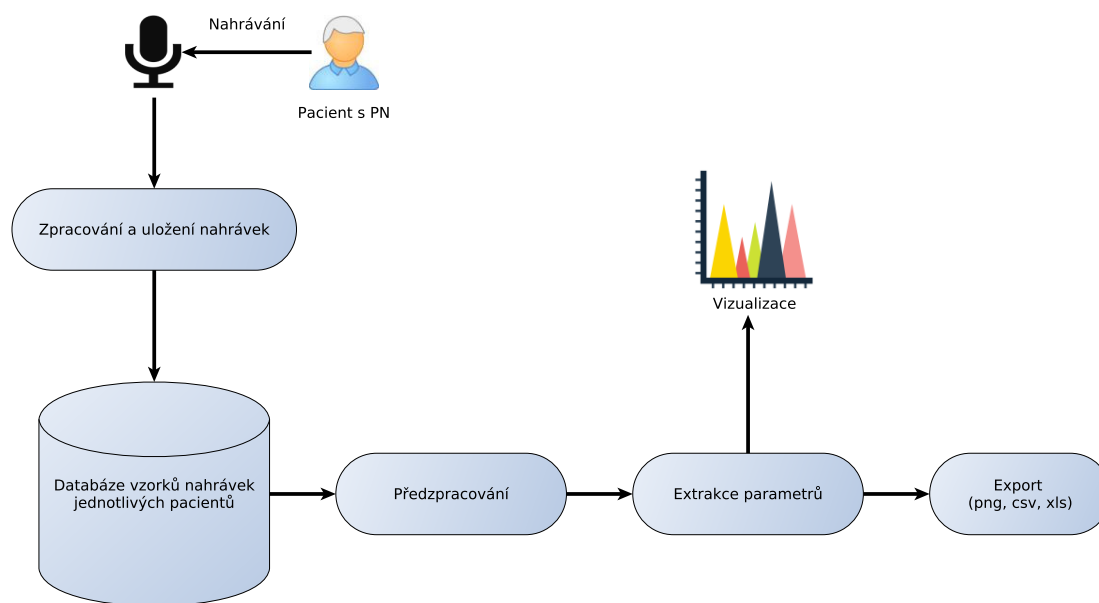
Celkově hlavní efekt byl nalezen pro celkové srovnání obyčejné a konverzační řeči a ukázalo se, že artikulační rychlost, $\text{mean}F_0$ a F_0 SD byly významně rozdílné v tomto srovnání.

4 NÁVRH SYSTÉMU ANALÝZY ŘEČI U PACIENTŮ S PARKINSONOVOU NEMOCÍ

V této kapitole je popsán systém pro analýzu záznamů řeči pacientů s Parkinsonovou nemocí v jejímž rámci bude výpočet mnoha různých parametrů používaných běžně v této analýze a součástí bude i základní statistická analýza těchto záznamů a výsledků.

4.1 Popis konceptu systému

Systém by měl být schopný vypočítat řečové parametry z vybraných nahrávek a dále vypočítat suprasegmentální parametry z vypočítaných řečových parametrů. Dále by měl být schopný tyto výsledky exportovat do některých standardních formátů a vizualizovat je.



Obr. 4.1: Schéma systému pro analýzu nahrávek

4.1.1 Test 3F: Dysartrický profil

Jelikož Hypokinetická dysartrie je souborem mnoha řečových poruch ať už ovlivňují řeč více či méně tak k vyhodnocení těchto poruch musíme v testu zahrnout mnoho komplexních řečových cvičení. Jedním takovým testem je *Test 3F: Dysartrický profil* což je klinický test využívaný pro mapování poruch řeči u získané dysartrie.

Byl vytvořen v roce 1977 pro český jazyk a v současnosti existuje již třetí revize, která je velmi často využívána v klinické praxi. Obsah byl od té doby zjednodušen a sloučen do tří částí: [16]

1. F1 faciokineze
2. F2 fonorespirace
3. F3 fonetiku

Výsledek testu je tzv. *dysartrický index DX*, který udává stupeň dysartrie a jehož hodnoty se pohybují v intervalu $< 0; 90 >$ kde stupnice vypadá takto:

- 0 - 16 velmi těžká dysartrie
- 17 - 35 těžká dysartrie
- 36 - 56 středně těžká dysartrie
- 57 - 74 lehká dysartrie
- 75 - 84 velmi lehká dysartrie
- 85 - 90 bez poruchy

Test zahrnuje mnoho menších úkolů, které mají za úkol zjistit stav hybnosti svalů potřebných k tvorbě řeči a kvalitu hlasové a řečové produkce.

4.1.2 Databáze nahrávek

K vytvoření řečového korpusu nebo-li databáze nahrávek řeči pacientů s Parkinsonovou nemocí slouží několik již existujících testovacích postupů, které se zaměřují na odhalení řečových vad a postižení. I přesto, že dysartrie je motorickou poruchou, její diagnostika závisí i na srozumitelnosti řeči což je také závislé na konkrétním jazyku a proto se tyto testy musí přizpůsobit konkrétnímu jazyku, v našem případě češtině.

Jako databáze nahrávek byla použita již existující databáze vytvořena na pracovišti I. Neurologické kliniky Lékařské fakulty Masarykovy univerzity a Fakultní nemocnice u sv. Anny v Brně, která nese název PARCZ (Czech Parkinsonian Speech Database). Tato databáze se zaznamenává podle navrženého protokolu a slouží k výzkumu řeči u pacientů s PN. V rámci výzkumu se plánuje také sledování progresu této nemoci realizovaným pomocí opakovaného nahrávání pacientů.

Databáze obsahuje nahrávky jak pacientů s Parkinsonovou nemocí tak zdravých lidí bez poruchy řeči pro porovnání. Pacienti jsou ještě před nahráváním klinicky vyšetřeni logopedem, neurologem a neurologickým psychologem. U každého bylo nahráno až 92 záznamů, i když ne u všech účastníků, které jsou seřazeny do složek jednotlivých účastníků. Složky jsou pojmenovány unikátním označením účastníka, které vypadají takto:

- **P1xxx** - žena s PN
- **P2xxx** - muž s PN

- **K1xxx** - žena bez PN
- **K2xxx** - muž bez PN

V tomto označení P a K nám říká zda se jedná o pacienta nebo kontrolního zdravého člověka. Následující číslo nám udává pohlaví (1 - žena, 2 - muž), a xxx je číslo označující daného člověka. V každé složce najdeme soubory nahrávek, které jsou ve formátu *wav* a pojmenovány ve stylu *ID_label_1.wav* kde *ID* je unikátní označení účastníka popsáno výše a *label* označuje danou úlohu například *9.2-1*, kdy mají účastníci za úkol přečíst text *Do čtvrt hodiny tam byla smršť*.

Poslední verze databáze PARCZ z data 1.10.2014 obsahuje nahrávky 101 pacientů s parkinsonovou nemocí (60 mužů a 41 žen) a 53 zdravých lidí (26 mužů a 27 žen) pro porovnání. Některé statistické údaje o věku účastníků jejichž nahrávky jsou v této databázi jsou uvedeny v tabulce 4.1.2.

Účastníci	Celkový počet		Průměrný věk		Minimální věk		Maximální věk	
	Muži	Ženy	Muži	Ženy	Muži	Ženy	Muži	Ženy
Zdraví	26	27	65.65	62.15	49	45	83	87
Pacienti	60	41	66.28	68.98	46	49	87	86

Tab. 4.1: Základní statistické údaje o databázi PARCZ

Databáze také obsahuje mnoho dalších ohodnocení účastníků jako například *dy-sartrický index*, *UPDRS* a další.

4.1.3 Předzpracování

Předzpracování slouží k úpravě zdrojových nahrávek nebo výpočtu některých údajů potřebných k samotnému výpočtu parametrů. Předzpracování zahrnuje operace jako *segmentace*, *převzorkování*, *normalizaci* aj.

4.1.4 Extrakce parametrů

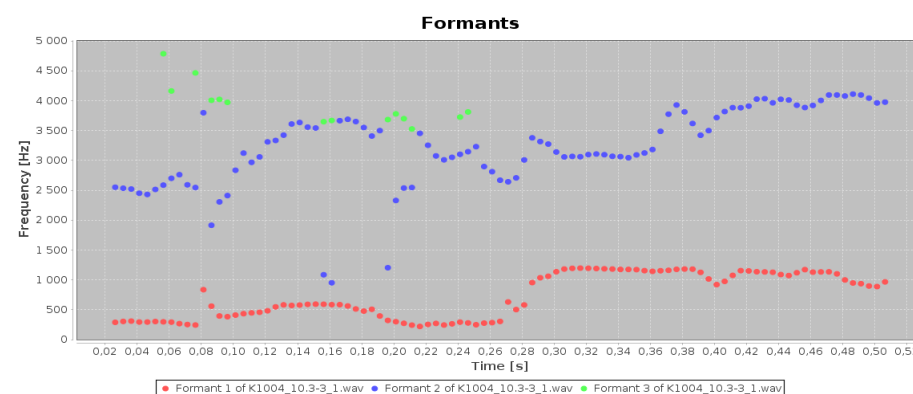
Extrakce parametrů zahrnuje samotný výpočet řečových parametrů na základě již předzpracovaných nahrávek a údajů získaných ve fázi předzpracování. Po samotné extrakci parametrů se dají vypočítat suprasegmentální parametry a nebo můžeme provést statistickou analýzu, což zahrnuje další výpočty, exportovat výsledky nebo je vizualizovat.

4.1.5 Export výsledků

Export probíhá po získání vypočítaných hodnot řečových parametrů. Tyto výsledky pak mohou být uloženy do různých formátů jako například *csv*.

4.1.6 Vizualizace

Vizualizace je fáze kdy můžeme výpočtené výsledky vykreslit například do grafů pro lepší znázornění a prezentaci výsledků. Vizualizace výsledků lze dále také exportovat například do formátu *png*. Příkladem vizualizace navržené aplikace zobrazující první tři formanty z jedné nahrávky je ukázána na 4.2



Obr. 4.2: Ukázka grafu prvních dvou formantů

5 POPIS SYSTÉMU

V této kapitole je detailněji popsán navrhovaný a implementovaný program pro výpočet řečových parametrů a technologie použité při jeho tvorbě. Součástí popisu jsou obrázky z aplikace zobrazující hlavní komponenty programu, popis hlavních funkcí a jeho použití. Jsou zde také popsány použité frameworky využívané při implementaci a jejich využití.

5.1 Použité technologie

Pro každou dílčí část systému byly použity specifické technologie k dosažení daných cílů. Tyto technologie jsou blíže popsány v následujících sekcích. Při výběru technologií byl brán zřetel na efektivitu a praktičnost využití technologií.

5.1.1 Aplikace

Samotná aplikace je implementována v jazyku *Java*, konkrétně ve verzi Java SE 1.8. Programovací jazyk Java je objektově orientovaný jazyk, který je v dnešní době velmi rozšířený díky své jednoduchosti, možnostem využití a nezávislosti na operačním systému [18]. Díky tomuto lze program spustit jak na systémech Windows tak na systémech typu UNIX nebo MAC.

Jako API (Application programming interface) pro grafické uživatelské rozhraní (GUI) byl použit framework *SWING*, který vychází ze starší verze *AWT*. Protože je tento framework kompletně napsán v Javě je multi-platformový a využívá architektury Model-View-Controller (MVC), která rozděluje program na *data*, *pohled* a *funkcionalitu* [19]. Tento framework umožňuje použití všech běžně využívaných komponent pro tvorbu *GUI*, díky kterým je pro uživatele snazší ovládání programu, jako například okna, tlačítka, tabulky atd. . . K vývoji byl použit vývojový program *Netbeans* k usnadnění a urychlení vývoje.

5.1.2 Výpočet řečových parametrů

K výpočtu většiny parametrů byl použit program *PRAAT*, který se specializuje na analýzu řeči. Tento program umožňuje také automatizaci výpočtů použitím vlastního skriptovacího jazyka. Je také možné jej spustit pomocí příkazového řádku což umožňuje jeho použití i skrze jiné programy. Těchto vlastností bylo využito v implementované aplikaci k výpočtu řečových parametrů.

Příklad skriptu programu praat

```
Read from file... 'file$'
object_name$ = selected$("Sound")
select Sound 'object_name$'

To Formant (burg)... 0.005 3 5500 0.025 50.0
table = Down to Table: "no", "yes", 6, "yes", 3, "yes", 3, "yes"

select Table 'object_name$'
Save as comma-separated file: outputFileFileName$
```

Praat z příkazové řádky

Program Praat umožňuje spouštět vlastní skripty přes příkazovou řádku a to následujícím způsobem: [12]

Windows Praat.exe --run script.praat parametr1 parametr2

Linux praat --run script.praat parametr1 parametr2

5.1.3 Vizualizace

K vizualizaci některých parametrů a k zobrazení některých dalších užitečných vlastností řeči byla použita knihovna zvaná *JFreeChart*, která je napsána v jazyku Java a poskytuje velké množství funkcionality k různému zobrazování dat. Umožňuje také export výsledných grafů do obrázkových formátů jako například *png* nebo *jpg*. Dalšími možnostmi vizualizovaných dat je interaktivní nastavení vzhledu grafů jako například zmenšení a zvětšení vybrané části grafu, změna barev os grafu a jiné.

5.1.4 Použité knihovny

Níže v tabulce 5.1.4 je uveden výčet java knihoven, jejich verze a využití, které byly použity při implementaci programu.

5.1.5 Náповěda programu

Pro realizaci nápovědy pro aplikaci bylo využito jazyků HTML [20] a CSS [21], které se používají pro implementaci statických internetových stránek. Jak již vyplývá z využití výše zmíněných jazyků, nápověda je implementována ve formě statických HTML stránek, které se dají zobrazit v programu.

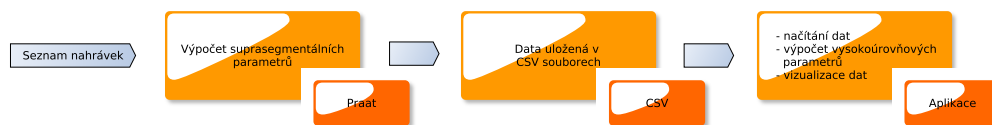
Název	Verze	Využití
log4j	2.7	logování programu
open-csv	2.2	práce s csv soubory
jfreechart	1.0.19	práce s grafy
commons-math3	3.3	matematické funkce použity při výpočtu
commons-io	2.5	funkce použity při práci se soubory

Tab. 5.1: Výčet knihoven

5.2 Postup výpočtu parametrů

V této sekci je detailní popis pro výpočet řečových parametrů. Popis zahrnuje ukázky z implementované aplikace a použití programu. Jsou zde popsány hlavní funkce programu a některé z jeho užitečných funkcí.

Postup výpočtu a jeho realizace je ukázána na obrázku 5.1 a níže najdete detailnější popis jednotlivých částí postupu.



Obr. 5.1: Postup výpočtu

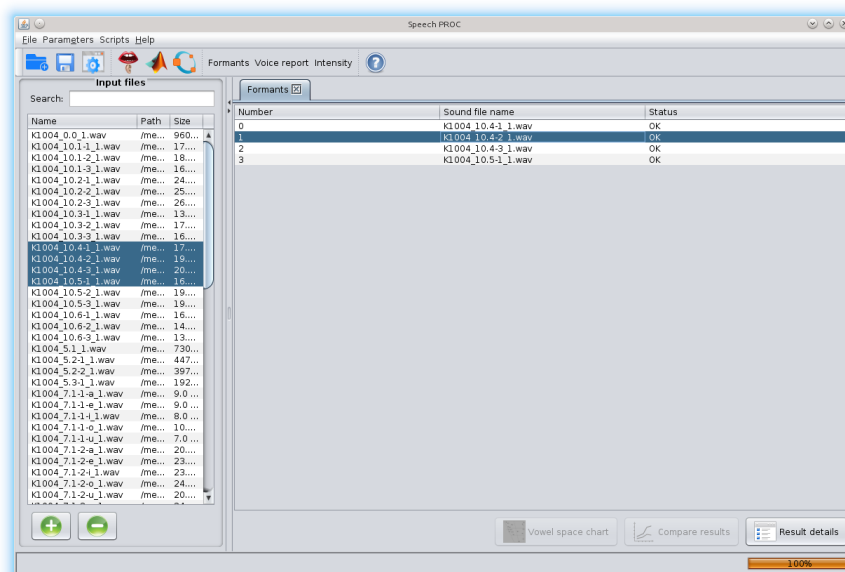
Blok Praat V této části se vypočítávají suprasegmentální parametry pomocí programu Praat pro vstupní soubory s nahrávkami. K těmto výpočtům se využívají praat skripty které se předávají pomocí volání programu přes příkazovou řádku.

Blok CSV Tento blok spojuje obě prostředí implementované aplikace a programu Praat. Výstupem výpočtů v programu Praat jsou csv soubory obsahující získané hodnoty suprasegmentálních parametrů.

Blok Aplikace Tyto soubory jsou pak načítány z aplikace a dále zpracovávány. Aplikace z těchto parametrů vypočítává vysokoúrovňové parametry, zobrazuje tyto data a poskytuje vizualizaci těchto dat.

5.3 Hlavní komponenty programu grafického uživatelského rozhraní

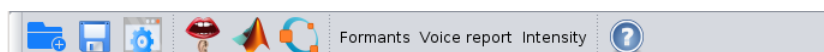
V této sekci jsou podrobně popsány hlavní komponenty aplikace a jejich účel, případně jejich využití. Celé hlavní okno aplikace je zobrazeno na obrázku 5.2. Všechny části hlavního okna jsou rozebrány v níže uvedených sekcích včetně jejich použití a funkcí.



Obr. 5.2: Hlavní okno aplikace





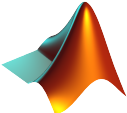


5.3.1 Nástrojová lišta

Hned pod hlavním menu programu se nachází nástrojová lišta obsahující tlačítka hlavních funkcí programu. Lišta je ukázána na obrázku 5.3 a funkce tlačítek v této liště je popsáno níže.



Obr. 5.3: Nástrojová lišta

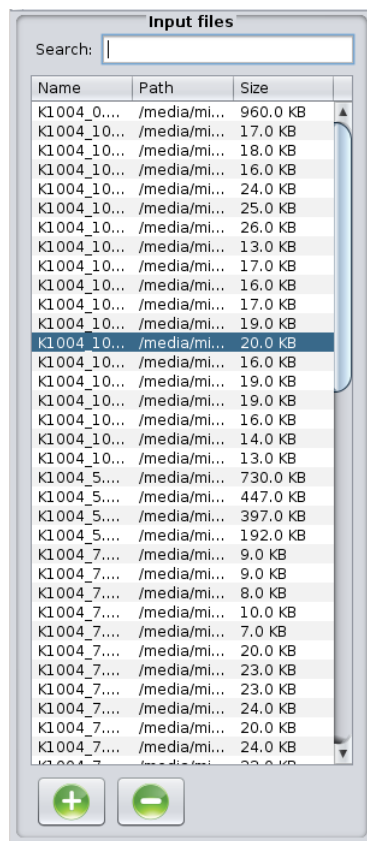
Níže uvedená tabulka 5.2 zobrazuje výčet tlačítek a jejich funkce v nástrojové liště.

Název	Ikona	Popis
Add files		Pro přidání zvukových souborů, které chceme analyzovat.
Save		Uložení vypočítaných dat ve formátu <i>zip</i>
Settings		Pro nastavení programu
Praat script		Spuštění Praat skriptu
Matlab script		Spuštění Matlab skriptu
Octave script		Spuštění Octave skriptu
Formants		Výpočet formantů z výbraných zvukových souborů
Voice report		Výpočet F0 a dalších parametrů z výbraných zvukových souborů
Intensity		Výpočet intenzity z výbraných zvukových souborů
Help		Zobrazení nápovědy

Tab. 5.2: Tlačítka v nástrojové liště a jejich ikony

5.3.2 Seznam zvukových souborů

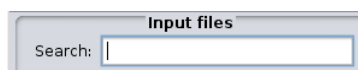
V levé části hlavního okna je zobrazen seznam zvukových souborů (viz. obrázek 5.4), které jsme si otevřeli v programu jako vstup pro výpočet řečových parametrů.



Obr. 5.4: Seznam zvukových souborů

Vyhledávání v seznamu

Nad seznamem zvukových nahrávek je pole pro vyhledávání v seznamu 5.5. Vyhledávání funguje tak, že vyhledává zadaný řetězec jako podřetězec v názvech souborů.



Obr. 5.5: Vyhledávací pole

Editace seznamu

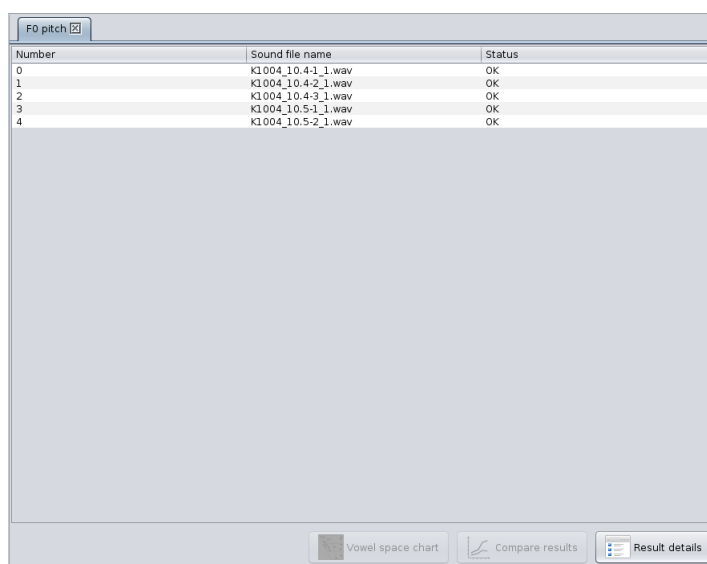
Pro editaci seznamu souborů slouží dvě tlačítka zobrazená pod tímto seznamem 5.6. Jedno tlačítko slouží pro přidání souborů a druhé pro odstranění souborů ze seznamu.



Obr. 5.6: Tlačítka pro editaci seznamu nahrávek

5.3.3 Hlavní panel s výsledky

Střed hlavního okna vyplňuje panel, ve kterém se zobrazuje seznam výsledků výpočtů pro jednotlivé soubory nahrávek, které uživatel vybral. Panel je zobrazen na obrázku k 5.7.



F0 pitch <input checked="" type="checkbox"/>		
Number	Sound file name	Status
0	K1004_10.4-1_1.wav	OK
1	K1004_10.4-2_1.wav	OK
2	K1004_10.4-3_1.wav	OK
3	K1004_10.5-1_1.wav	OK
4	K1004_10.5-2_1.wav	OK

Below the table, there are three buttons: "Vowel space chart", "Compare results", and "Result details".

Obr. 5.7: Okno s výsledky

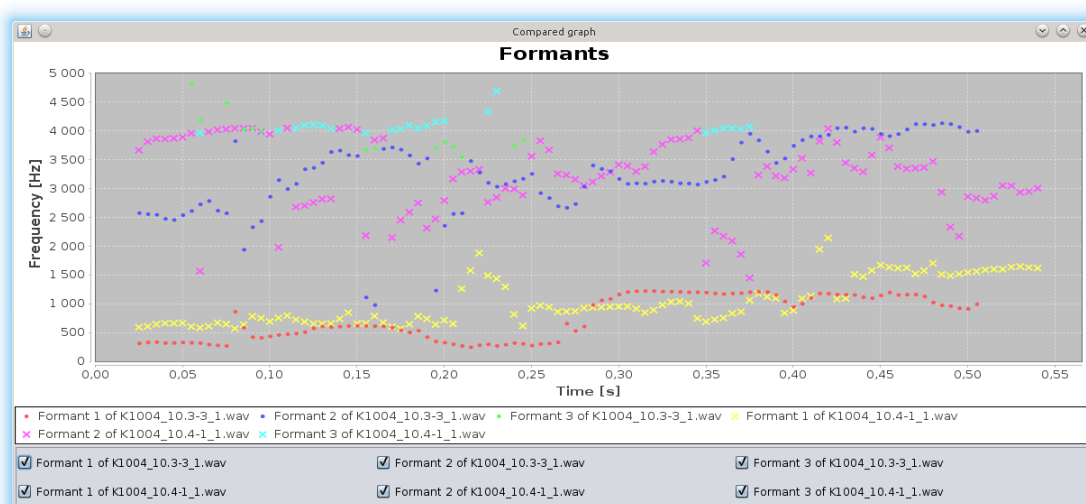
5.3.4 Okno s hodnotami výpočtů

Pokud je výpočet zvolených parametrů úspěšně ukončen uživatel si může zobrazit výsledné hodnoty řečových parametrů zvukové nahrávky dvojklikem na příslušný řádek v tabulce výsledků a nebo tlačítkem *Result details* na spodu panelu 5.3.3 jak již bylo popsáno výše. Případně lze tyto výsledky zobrazit i klávesou *ENTER* po výběru řádku. Uživatel si může stejnou cestou zobrazit výsledky více souborů současně pokud vybere více řádků v seznamu výsledků.

5.3.5 Visualizace výsledků

Graf vypočítaných hodnot si uživatel může zobrazit pomocí tlačítka *Show chart* v okně s detaily výsledků, které je popsáno výše v sekci 5.3.4. Po máčknutí tohoto tlačítka se zobrazí nové okno s grafem pro dané výsledky. V tomto novém okně si uživatel také může interaktivně vybírat které výsledky chce zobrazovat v grafu. Tato funkce je velmi užitečná u grafů s mnoha hodnotami, případně u porovnávacích grafů, které se mohou stát nepřehlednými kvůli zobrazování příliš mnoho dat, které uživatel nemusí v danou chvíli potřebovat. Konkrétní grafy pro konkrétní parametry jsou zobrazeny v následující sekci 5.4.

Program umožňuje i grafické porovnání výsledků ze 2 zvukových nahrávek. Stačí označit v panelu s výsledky 2 řádky výsledků, které chceme porovnat a zmáčknout tlačítko *Compare results*. Poté se zobrazí okno s grafem ve kterém budou zobrazeny hodnoty výsledků získané z obou nahrávek. Uživatel si samozřejmě opět může interaktivně volit, které data chce zobrazovat. Příkladem takového porovnání je vidět na obrázku 5.8, kde jsou porovnány formanty získané ze dvou nahrávek.

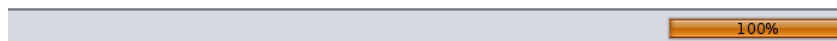


Obr. 5.8: Grafické porovnání 2 výsledků

Zobrazený graf je možné libovolně zvětšovat a zmenšovat případně nastavit některé další možnosti jako barvu pozadí, font popisu hlavních os a jiné. Všechny tyto možnosti lze dosáhnout pomocí vyskakovacího menu po stisknutí pravého tlačítka myši na ploše grafu. Přes toto menu lze i exportovat graf do formátu *png*.

5.3.6 Stavová lišta

Na spodu hlavního okna je stavová lišta zobrazující průběh výpočtu. Toto je velmi užitečné pokud si uživatel zvolí výpočet řečových parametrů z mnoha zvukových nahrávek. Stavovou lištu můžeme vidět na obrázku 5.9.



Obr. 5.9: Stavová lišta

5.4 Hlavní funkce programu

V této sekci jsou popsány implementované hlavní funkce programu a jeho možnosti, které poskytuje uživateli. Mezi tyto funkce hlavně patří výpočet řečových parametrů, kvůli kterému byla aplikace implementována, a zobrazení výsledků ať už vypočítaných hodnot nebo grafické znázornění těchto hodnot. Konkrétní ukázky zmíněných funkcí jsou uvedeny v následujících sekcích.

5.4.1 Načtení zvukových souborů

Pro výpočet parametrů je nutné zvolit zvukové nahrávky ze kterých se tyto parametry získají. Uživatel si tak sám může zvolit soubory formátu *.wav*, na základě kterých se bude výpočet provádět. Toto se v programu provádí ve dvou krocích. V prvním kroku si uživatel vybere zvukové soubory, které má uloženy pevném úložišti (např. pevný disk počítače na kterém aplikaci spustil) a aplikace si načte informace o těchto souborech a zobrazí seznam v levé části panelu, který je popsán v sekci 5.3.2. V druhém kroku si uživatel z tohoto seznamu může vybrat libovolný počet souborů, na základě kterých bude aplikace provádět výpočet. Tento seznam si uživatel může samozřejmě interaktivně upravovat.

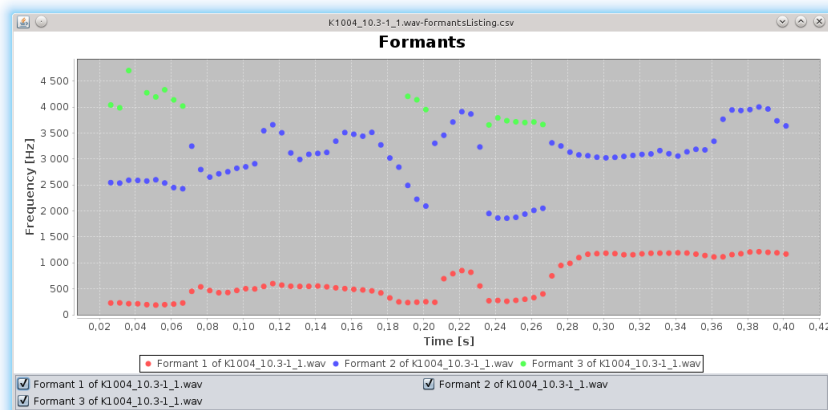
5.4.2 Výpočet formantů

Jednou z hlavních funkcí aplikace je výpočet formantů z vybraných nahrávek. Tohoto může uživatel dosáhnout několika cestami v grafickém rozhraní aplikace. Jednou z cest je položka v hlavním menu programu, další možností je tlačítko v nástrojové liště (viz. 5.3.1) a nebo klávesovou zkratkou. Po skončení výpočtu se vytvoří nový panel v hlavním panelu s výsledky (viz. 5.3.3) ve kterém se zobrazí stavy výpočtů pro jednotlivé soubory. Uživatel si pak tyto výsledky může dále zobrazit jak je popsáno níže. K výpočtu si uživatel musí také zvolit vstupní parametry použité při výpočtu.

Formants of K1004_10.3-1_1.wav									
time(s)	intensity	informants	F1(Hz)	B1(Hz)	F2(Hz)	B2(Hz)	F3(Hz)	B3(Hz)	
0.026281	0.0000005	3	228.649	120.511	2545.649	461.351	4038.477	955.344	
0.031281	0.0000005	3	230.100	114.548	2535.381	468.915	3984.284	1366.953	
0.036281	0.0000005	3	214.931	115.730	2590.935	271.293	4701.663	2451.324	
0.041281	0.0000004	2	211.479	138.163	2587.444	314.842	—undefined—	—undefined—	
0.046281	0.0000004	3	195.744	148.882	2576.066	410.277	4275.765	1805.148	
0.051281	0.000001	3	186.610	167.898	2599.446	368.371	4193.168	1512.731	
0.056281	0.000002	3	195.317	166.400	2538.238	367.287	4351.034	1528.015	
0.061281	0.000002	3	207.387	159.700	2448.967	380.923	4138.083	1126.932	
0.066281	0.000002	3	227.951	205.849	2426.687	651.198	4017.506	1047.078	
0.071281	0.000002	2	451.279	561.308	3246.109	1390.904	—undefined—	—undefined—	
0.076281	0.000003	2	538.279	314.454	2796.053	1015.542	—undefined—	—undefined—	
0.081281	0.000004	2	466.322	193.783	2651.869	638.711	—undefined—	—undefined—	
0.086281	0.000005	2	425.456	185.323	2714.588	609.151	—undefined—	—undefined—	
0.091281	0.000006	2	427.825	200.496	2753.820	597.981	—undefined—	—undefined—	
0.096281	0.000006	2	468.863	267.592	2821.836	682.951	—undefined—	—undefined—	
0.101281	0.000006	2	501.194	286.518	2848.622	612.583	—undefined—	—undefined—	
0.106281	0.000006	2	497.983	246.701	2908.202	628.722	—undefined—	—undefined—	
0.111281	0.000006	2	547.369	286.675	3543.475	1204.437	—undefined—	—undefined—	
0.116281	0.000006	2	600.998	315.327	3658.867	938.898	—undefined—	—undefined—	
0.121281	0.000006	2	573.030	267.756	3504.687	1238.548	—undefined—	—undefined—	
0.126281	0.000006	2	550.078	207.983	3116.119	1304.674	—undefined—	—undefined—	
0.131281	0.000006	2	547.022	187.896	2991.529	1012.654	—undefined—	—undefined—	
0.136281	0.000006	2	548.775	181.625	3088.365	1083.120	—undefined—	—undefined—	
0.141281	0.000006	2	553.659	182.287	3106.264	1035.649	—undefined—	—undefined—	
0.146281	0.000006	2	539.455	159.093	3126.468	1110.281	—undefined—	—undefined—	

Formant 1	Formant 2	Formant 3	Bandwidth 1	Bandwidth 2	Bandwidth 3
Mean: 673.2810263157895 Hz	Mean: 3029.643736842105 Hz	Mean: 3997.2380555555555 Hz	Mean: 368.3711710526317 Hz	Mean: 904.1404342105265 Hz	Mean: 1247.7425555555555 Hz
Median: 547.1904999999999 Hz	Median: 3088.2185 Hz	Median: 4000.895 Hz	Median: 243.184 Hz	Median: 761.4205 Hz	Median: 1190.248 Hz
Stdev: 377.8192493826274 Hz	Stdev: 551.8678083462602 Hz	Stdev: 278.7010685691893 Hz	Stdev: 314.15344426776494 Hz	Stdev: 615.3404429497446 Hz	Stdev: 471.1264023274695 Hz
Variance: 142747.38520405197 Hz	Variance: 304558.0778889046 Hz	Variance: 77674.28562160797 Hz	Variance: 98692.3895452997 Hz	Variance: 378643.86072958785 Hz	Variance: 221960.08697002468 Hz
Min: 186.61 Hz	Min: 1858.436 Hz	Min: 3655.962 Hz	Min: 89.645 Hz	Min: 163.66 Hz	Min: 702.94 Hz
Max: 1214.991 Hz	Max: 4001.432 Hz	Max: 4701.663 Hz	Max: 1784.774 Hz	Max: 3617.505 Hz	Max: 2451.324 Hz
1.p: 186.61 Hz	1.p: 1858.436 Hz	1.p: 3655.962 Hz	1.p: 89.645 Hz	1.p: 163.66 Hz	1.p: 702.94 Hz
99.p: 186.61 Hz	99.p: 1858.436 Hz	99.p: 3655.962 Hz	99.p: 89.645 Hz	99.p: 163.66 Hz	99.p: 702.94 Hz
IR: 0.0 Hz	IR: 0.0 Hz	IR: 0.0 Hz	IR: 0.0 Hz	IR: 0.0 Hz	IR: 0.0 Hz

Obr. 5.10: Detailní výsledky výpočtu formantů



Obr. 5.11: Graf hodnot formantů

Zobrazení vypočítaných hodnot

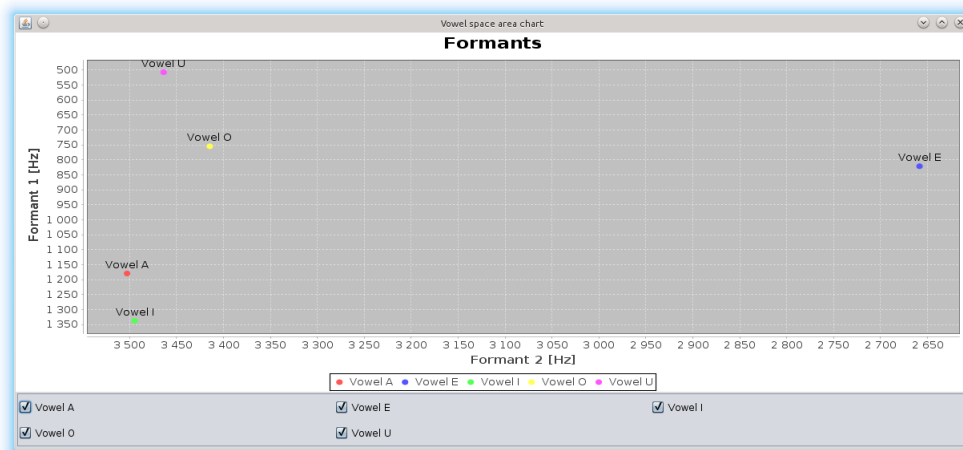
Zobrazení výsledných hodnot formantů a šířek pásma daných formantů se zobrazuje v tabulce v novém okně. Vysokoúrovňové parametry jsou vypočítány dodatečně a vypsány ve spodním panelu zobrazeného okna. Ukázka tohoto okna je na obrázku 5.10

Vizualizace výsledků

Výsledné hodnoty formantů si může uživatel zobrazit v grafickém zobrazení v podobě grafu. Příklad tohoto grafu můžeme vidět na obrázku 5.11.

Graf pozice samohlásek

Jedním z užitečných grafických zobrazení je zobrazení pozice formantů samohlásek. Tohoto docílí uživatel když si nechá programem vypočítat formanty 5-ti samohlásek a v panelu výsledků zmáčkne tlačítko „Vowel space chart“. Zobrazí se graf kde na ose x je hodnota prvního formantu dané samohlásky a na ose y je hodnota druhého formantu. Příklad tohoto grafu je na obrázku 5.12.



Obr. 5.12: Graf pozic samohlásek

5.4.3 Výpočet F_0

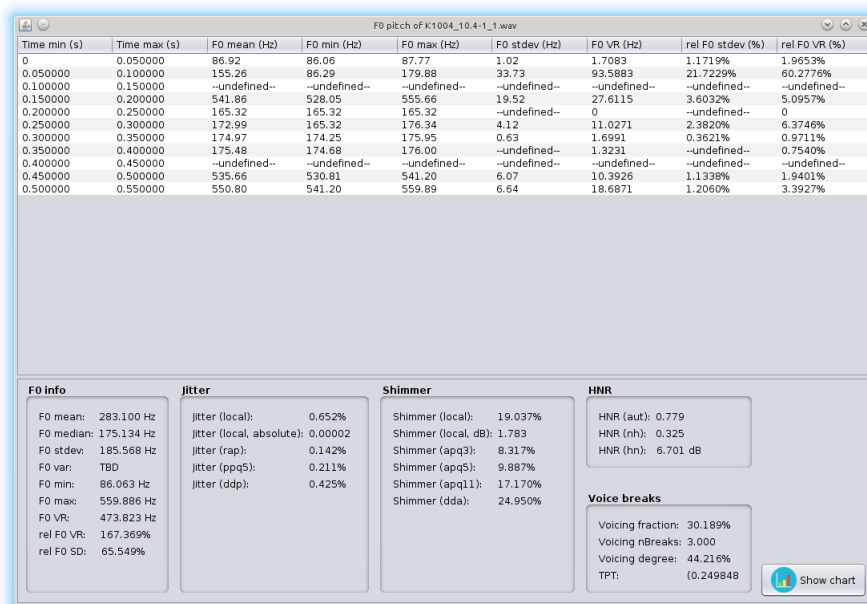
Další hlavní funkcí je výpočet základní frekvence tónu F_0 . V programu je výpočet F_0 zahrnut ve funkci s názvem *Voice report*. Tato funkce nejen vypočítává F_0 ale i další řečové parametry jako *Jitter*, *Shimmer*, některé pauzovací parametry a další. K výpočtu si uživatel opět musí zvolit vstupní parametry použité při výpočtu F_0 .

Zobrazení vypočítaných hodnot

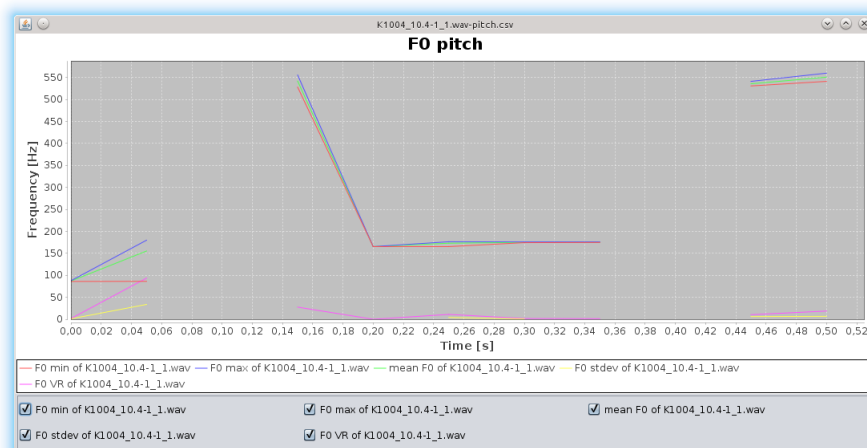
Vypočítané parametry pomocí funkce *Voice report* v programu jsou zobrazeny v okně výsledků, kde parametry *Jitter* a *Shimmer* jsou zobrazeny ve spodním panelu tohoto okna současně s některými pauzovacími parametry a vysokoúrovňovými parametry F_0 . Okno je zobrazeno na obrázku 5.13

Vizualizace výsledků

Zobrazení grafu uživatel opět provede stisknutím tlačítka *Show chart*, ve kterém jsou zobrazeny hodnoty fundamentální frekvence. Ukázka tohoto grafu lze vidět na obrázku 5.14.



Obr. 5.13: Detailní výsledky funkce *Voice report*



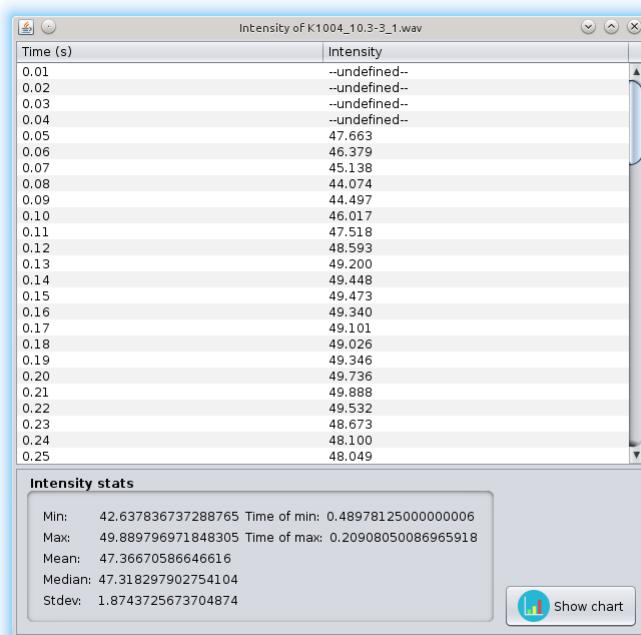
Obr. 5.14: Graf hodnot F0

5.4.4 Výpočet intensity

Poslední funkcí sloužící k výpočtu řečových parametrů je výpočet intensity. Hodnoty intensity jsou vypočítávány pro daný čas s časovým krokem 0.01 s.

Zobrazení vypočítaných hodnot

Příklad zobrazení výsledků intensity a její vysokoúrovňové parametry jsou ukázány na obrázku 5.15.



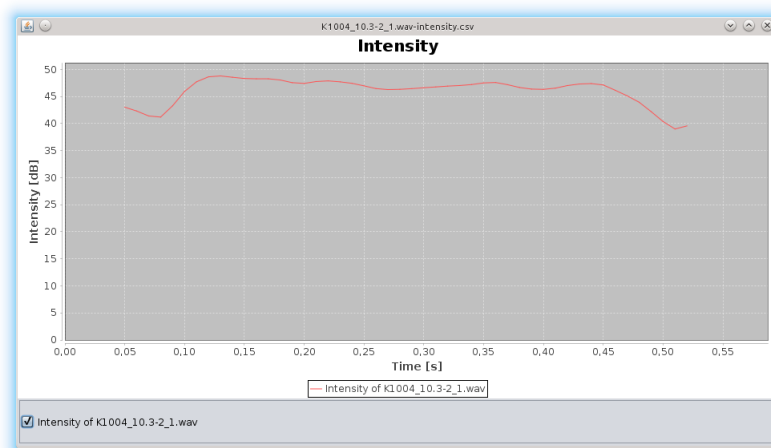
Obr. 5.15: Detailní výsledky hodnot intensity

Vizualizace výsledků

Příkladem grafického zobrazení intensity je obrázek 5.16

5.4.5 Spouštění externích skriptů

V rámci implementované aplikace lze také spouštět skripty programů *Praat*, *Matlab* nebo *Octave*. Ovšem ke správnému fungování těchto funkcí je nutné mít správně nastaveny cesty k těmto programům v nastavení programu (viz. 5.4.6).



Obr. 5.16: Graf hodnot intenzity

Praat

Kromě toho, že samotný program využívá k výpočtu řečových parametrů program *Praat* může také pouštět vlastní skripty uživatele. Využívá k tomu volání programu *Praat* z příkazové řádky. Uživatel si pouze vybere soubor obsahující skript a program jej spustí.

Matlab

Dalším typem skriptů, které je program schopen spustit jsou skripty pro program *Matlab*. Opět se využívá volání programu z příkazové řádky a cesta k programu *Matlab* musí být správně nastavena v nastavení programu.

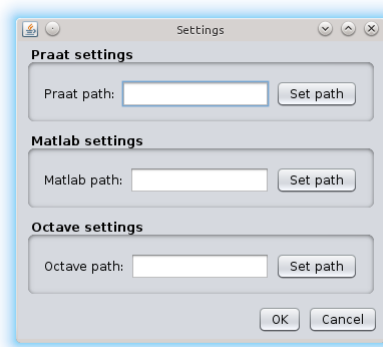
Octave

Stejně jako předchozí typy skriptů i skripty pro program *Octave* se dají spustit pomocí příkazové řádky a opět musí být cesta správně nastavena.

5.4.6 Nastavení programu

Pro správné fungování všech funkcí programu je potřeba nastavit cesty k programům *Praat*, *Matlab* a *Octave*. Toto nastavení se provádí pomocí okna nastavení 5.17, které se otevře pomocí položky v hlavním menu *Settings* nebo pomocí tlačítka v liště nástrojů.

Nastavení cesty k programu *Praat* je nutné pro správné fungování výpočtů řečových parametrů.



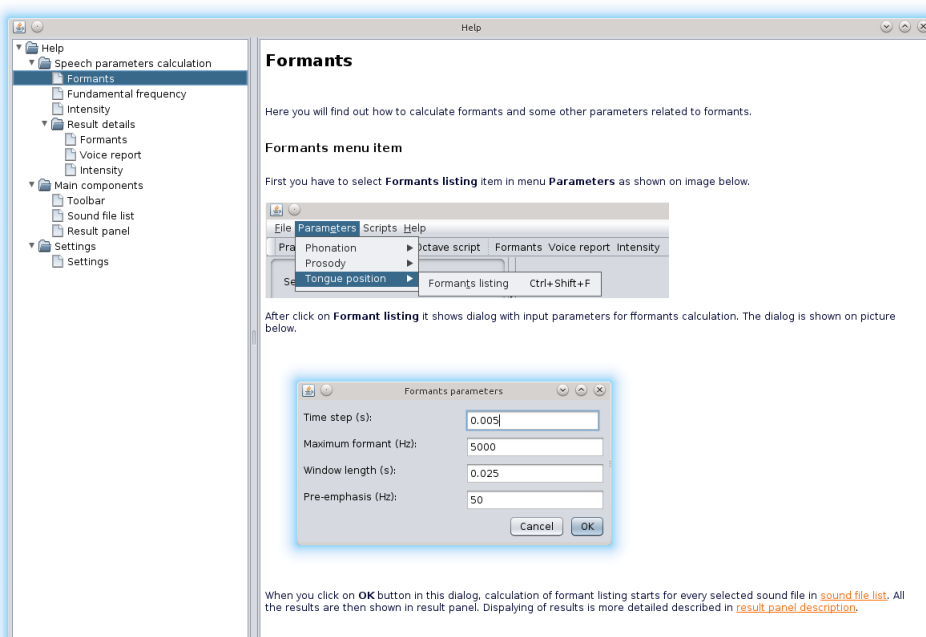
Obr. 5.17: Nastavení programu

5.4.7 Náповěda

Program také obsahuje svoji vlastní nápovědu, která lze otevřít pomocí klávesové zkratky *F1* nebo v hlavním menu *Help* a nebo tlačítkem v nástrojové liště. Po oteření nápovědy se zobrazí okno s obsahem v jeho levé části a popisem vybrané sekce v pravé části. Stránky nápovědy jsou ve formě statických internetových stránek HTML. Okno s nápovědou je zobrazeno na obrázku 5.18.

5.4.8 Logování programu

Implementovaný program ukládá do logu většinu akcí co uživatel pomocí programu udělá a akce co provádí, případně chyby které nastanou. Tyto informace jsou všechny uloženy ve složce *logs* v adresáři, ze kterého byl program spuštěn. K logování se využívá java knihovny *log4j* a její *xml* konfigurace.



Obr. 5.18: Nápořěda programu

6 SOUHRN VYPOČÍTANÝCH PARAMETRŮ

V této kapitole je výčet všech řečových parametrů, které je program schopen vypočítat, jejich jednotky a popis významu. Tyto parametry jsou rozděleny do kategorií podle efektu na řeč.

Pořadí	Název	Jednotka	Popis
1	F_0 median	[Hz]	medián F_0
2	F_0 mean	[Hz]	střední hodnota F_0
3	F_0 stdev	[Hz]	směrodatná odchylka F_0
4	F_0 var	[Hz]	rozptyl F_0
5	F_0 max	[Hz]	maximální hodnota F_0
6	F_0 min	[Hz]	minimální hodnota F_0
7	F_0 VR	[Hz]	variační rozsah F_0
8	rel F_0 SD	[%]	relativní kolísání F_0
9	rel F_0 VR	[%]	relativní směrodatná odchylka F_0
10	jitter(local)	[%]	lokální kolísání periody
11	jitter(local,abs)	[s]	absolutní hodnota lokální kolísání periody
12	jitter(rap)	[%]	průměrný rozdíl mezi periodou a průměrem dvou sousedních period
13	jitter(ppq5)	[%]	průměrný rozdíl mezi periodou a průměrem čtyř sousedních period
14	jitter(ddp)	[%]	průměrný rozdíl mezi za sebou jdoucími rozdíly period
15	shimmer(local)	[%]	lokální kolísání amplitudy
16	shimmer(local,dB)	[dB]	lokální kolísání amplitudy v dB
17	shimmer(apq3)	[%]	průměrný rozdíl mezi amplitudou a průměrem dvou sousedních amplitud
18	shimmer(apq5)	[%]	průměrný rozdíl mezi amplitudou a průměrem čtyř sousedních amplitud
19	shimmer(apq11)	[%]	průměrný rozdíl mezi amplitudou a průměrem deseti sousedních amplitud
20	Intensity	[dB]	intenzita zvuku v daném čase

21	Intensity min	[dB]	minimální hodnota intenzity zvuku
22	Intensity max	[dB]	maximální hodnota intenzity zvuku
23	Intensity median	[dB]	medián intenzity zvuku
24	Intensity mean	[dB]	střední hodnota intenzity zvuku
25	Intensity stdev	[dB]	směrodatná odchylka intenzity zvuku

Tab. 6.1: Parametry popisující fonaci

Pořadí	Název	Jednotka	Popis
26	TPT	[s]	celkové trvání pauz delších jak 10 ms
27	voicing fraction	[%]	shluk neznělých úseku
28	voicing nbreaks	[-]	počet pauz
29	voicing degree	[%]	počet pauz v %

Tab. 6.2: Parametry popisující pauzování

Pořadí	Název	Jednotka	Popis
30	F_1 median	[Hz]	medián F_1
31	F_1 mean	[Hz]	střední hodnota F_1
32	F_1 stdev	[Hz]	směrodatná odchylka F_1
33	F_1 var	[Hz]	rozptyl F_1
34	F_1 max	[Hz]	maximální hodnota F_1
35	F_1 min	[Hz]	minimální hodnota F_1
36	F_1 1.p	[Hz]	1.percentil F_1
37	F_1 99.p	[Hz]	99.percentil F_1
38	F_1 ir	[Hz]	mezipercentilový rozsah F_1
39	B_1 median	[Hz]	medián B_1
40	B_1 mean	[Hz]	střední hodnota B_1
41	B_1 stdev	[Hz]	směrodatná odchylka B_1
42	B_1 var	[Hz]	rozptyl B_1
43	B_1 max	[Hz]	maximální hodnota B_1
44	B_1 min	[Hz]	minimální hodnota B_1
45	B_1 1.p	[Hz]	1.percentil B_1
46	B_1 99.p	[Hz]	99.percentil B_1

47	B_1 ir	[Hz]	mezipercentilový rozsah B_1
48	F_2 median	[Hz]	medián F_2
49	F_2 mean	[Hz]	střední hodnota F_2
50	F_2 stdev	[Hz]	směrodatná odchylka F_2
51	F_2 var	[Hz]	rozptyl F_2
52	F_2 max	[Hz]	maximální hodnota F_2
53	F_2 min	[Hz]	minimální hodnota F_2
54	F_2 1.p	[Hz]	1.percentil F_2
55	F_2 99.p	[Hz]	99.percentil F_2
56	F_2 ir	[Hz]	mezipercentilový rozsah F_2
57	B_2 median	[Hz]	medián B_2
58	B_2 mean	[Hz]	střední hodnota B_2
59	B_2 stdev	[Hz]	směrodatná odchylka B_2
60	B_2 var	[Hz]	rozptyl B_2
61	B_2 max	[Hz]	maximální hodnota B_2
62	B_2 min	[Hz]	minimální hodnota B_2
63	B_2 1.p	[Hz]	1.percentil B_2
64	B_2 99.p	[Hz]	99.percentil B_2
65	B_2 ir	[Hz]	mezipercentilový rozsah B_2
66	F_3 median	[Hz]	medián F_3
67	F_3 mean	[Hz]	střední hodnota F_3
68	F_3 stdev	[Hz]	směrodatná odchylka F_3
69	F_3 var	[Hz]	rozptyl F_3
70	F_3 max	[Hz]	maximální hodnota F_3
71	F_3 min	[Hz]	minimální hodnota F_3
72	F_3 1.p	[Hz]	1.percentil F_3
73	F_3 99.p	[Hz]	99.percentil F_3
74	F_3 ir	[Hz]	mezipercentilový rozsah F_3
75	B_3 median	[Hz]	medián B_3
76	B_3 mean	[Hz]	střední hodnota B_3
77	B_3 stdev	[Hz]	směrodatná odchylka B_3
78	B_3 var	[Hz]	rozptyl B_3
79	B_3 max	[Hz]	maximální hodnota B_3
80	B_3 min	[Hz]	minimální hodnota B_3
81	B_3 1.p	[Hz]	1.percentil B_3
82	B_3 99.p	[Hz]	99.percentil B_3

83	B_3 ir	[Hz]	mezipercentilový rozsah B_3
----	----------	------	-------------------------------

Tab. 6.3: Parametry popisující hybnost jazyka

Pořadí	Název	Jednotka	Popis
84	HNr(aut)	[-]	střední hodnota autokolerace
85	HNr(nh)	[-]	střední hodnota poměru signál- šum
86	HNr(hn)	[dB]	střední hodnota poměru signál- šum v dB

Tab. 6.4: Parametry popisující kvalitu řeči

7 ZÁVĚR

Tato práce se zabývá Parkinsonovou nemocí a Hypokinetickou dysartrií a analýzou řeči při jejich postižení. V první kapitole je představena Parkinsonova nemoc a Hypokinetická dysartrie. Také se zabývá popisem oblastí fonace, prozodie, artikulace a plynulost řeči a jak jsou tyto oblasti řeči ovlivněny při výskytu Hypokinetické dysartrie.

Druhá kapitola se již zabývá výčtem a popisem některých nejčastěji používaných řečových parametrů při analýze. Tyto parametry popisují Hypokinetickou dysartrii mezi něž patří frekvence základního tónu, první a druhý formant, různé pauzovací parametry, parametry zaměřené na tempo řeči, intenzitu aj.

V další kapitole jsou popsány některé významné studie zaměřené na analýzu řeči u lidí s Parkinsonovou nemocí. Těchto studií se dobrovolně účastnili pacienti a zdraví lidé pro porovnání získaných výsledků. Autoři pomocí vypočítaných parametrů analyzovali řeč pacientů a zjišťovali význam těchto parametrů u Hypokinetické dysartrie spojené s Parkinsonovou nemocí. U každé ze zmíněných studií je popsán její cíl, účastníci a jejich popis, průběh studie a její výsledky a zjištění.

Ve čtvrté kapitole je popsán návrh systému na výpočet řečových parametrů a základní statistické analýzy, který je v rámci této práce implementován. Je zde popsáno schéma tohoto systému a jeho dílčí části a jejich funkce. Kapitola také obsahuje popis databáze nahrávek řeči pacientů a kontrolních lidí, která se použila jako zdroj pro výpočet řečových parametrů. Také byl zmíněn postup vzniku této databáze, konkrétně Test 3F, hodnotící stav dysartrie.

V kapitole číslo 5 je popsána implementace a realizace tohoto systému. Na začátku kapitoly lze nalézt výčet a popis použitých technologií při implementaci aplikace a výpočtu řečových parametrů. Součástí kapitoly je i popis hlavních částí grafického uživatelského rozhraní aplikace včetně ukázek z aplikace a v neposlední řadě zde lze nalézt detailní popis hlavních funkcí aplikace a jejich názorné ukázky.

V poslední kapitole je výčet všech řečových parametrů, které je schopen výsledný program vypočítat jako výsledek práce. Parametry jsou řezeny podle toho jak ovlivňují řeč.

Jelikož byl na implementaci této aplikace omezený čas byla navržena tak aby se dala jednoduše rozšiřovat o další funkce jako například výpočet dalších řečových parametrů, možnosti dodatečného grafického zobrazení vypočítaných dat v podobě grafů a dalších užitečných funkcí. V rámci tohoto rozšiřování byla již implementována možnost spouštět skripty i pro jiné programy jako je například *Matlab*.

LITERATURA

- [1] MEKYSKA, J.; REKTOROVA, I.; SMEKAL, Z. *Selection of Optimal Parameters for Automatic Analysis of Speech Disorders in Parkinson's Disease* [online]. Telecommunications and Signal Processing (TSP), 2011 34th International Conference on, 2011: s. 408–412. [cit. 25. 11. 2016]. Dostupné z URL: <http://dx.doi.org/10.1109/TSP.2011.6043700>.
- [2] ZAMIŠKOVÁ, G.; RESSNER, P.; DLOUHÁ, J.; ŠIGUTOVÁ, D. *Poruchy řeči u Parkinsonovy nemoci* [online]. Neurologie pro Praxi, ročník 11, č. 2, 2010: s. 112–116 [cit. 17. 11. 2014]. Dostupné z URL: <http://www.neurologiepropraxi.cz/pdfs/neu/2010/02/10.pdf>
- [3] OROZCO-ARROYAVE, J.R.; ARIAS-LONDOÑO, J.D.; VARGAS-BONILLA, J.F. *Analysis of Speech from People with Parkinson's Disease through Nonlinear Dynamics*.
- [4] MEKYSKA, J.; SMĚKAL, Z.; KOŠTÁLOVÁ, M.; MRAČKOVÁ, M.; SKUTILOVÁ S.; REKTOROVÁ I. *Motorické aspekty poruch řeči u Parkinsonovy nemoci a jejich hodnocení* [online]. 2014, Česká a Slovenská neurologie a neurochirurgie, 2011. Dostupné z URL: <http://www.csnn.eu/ceska-slovenska-neurologie-clanek/motoricke-aspekty-poruch- reci-u-parkinsonovy-nemoci-a-jejich-hodnoceni-36308>.
- [5] MASSANO, J.; BHATIA, K. P. *Clinical Approach to Parkinson's Disease: Features, Diagnosis, and Principles of Management* [online]. Cold Spring Harb Perspect Med, ročník 2, č. 6, 2012. Dostupné z URL: <https://www.ncbi.nlm.nih.gov/pubmed/22675666>.
- [6] GOBERMAN, A. M.; COELHO, C. *Acoustic analysis of parkinsonian speech I: speech characteristics and L-Dopa therapy* [online]. NeuroRehabilitation, ročník 17, č. 3, 2002: s. 237–246 [cit. 19. 11. 2016]. Dostupné z URL: <http://www.ncbi.nlm.nih.gov/pubmed/12237505>
- [7] MEKYSKA, J.; REKTOROVA, I.; SMEKAL, Z. *Selection of Optimal Parameters for Automatic Analysis of Speech Disorders in Parkinson's Disease* [online]. Telecommunications and Signal Processing (TSP), 2011 34th International Conference on, 2011: s. 408–412. [cit. 19. 11. 2016]. Dostupné z URL: <http://dx.doi.org/10.1109/TSP.2011.6043700>

- [8] SKODDA, S.; GRÖNHEIT, W.; MANCINELLI, N.; SCHLEGEL, U. *Progression of Voice and Speech Impairment in the Course of Parkinson's Disease: A Longitudinal Study* [online]. Parkinson's Disease, ročník 2013, Article ID 389195, 2013: s. 1-8. [cit. 20. 11. 2016]. Dostupné z URL: <http://dx.doi.org/10.1155/2013/389195>.
- [9] SKODDA, S.; RINSCHKE, H.; SCHLEGEL, U. *Progression of dysprosody in Parkinson's disease over time – a longitudinal study* [online]. Movement Disorders, ročník 24, č. 5, 2009: s. 716-722. [cit. 20. 11. 2016]. Dostupné z URL: <http://www.ncbi.nlm.nih.gov/pubmed/19117364>.
- [10] SKODDA, S.; SCHLEGEL, U. *Speech Rate and Rhythm in Parkinson's Disease* [online]. Movement Disorders, ročník 23, č. 7, 2008: s. 985-992. [cit. 20. 11. 2016]. Dostupné z URL: https://www.researchgate.net/publication/5469820_Speech_rate_and_rhythm_in_Parkinson's_Disease
- [11] RUSZ, J.; CMEJLA, R.; RUZICKOVA, H.; RUZICKA, E. *Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease* [online]. J Acoust Soc Am, ročník 129, č. 1, 2011: s. 350–367. [cit. 21. 11. 2016]. Dostupné z URL: <http://www.ncbi.nlm.nih.gov/pubmed/21303016>
- [12] BOERSMA, P.; WEENIK, D. *Praat: doing phonetics by computer [Computer program]* [online]. Version 5.3.51, retrieved 2 June 2013 [cit. 22. 11. 2016]. Dostupné z URL: <http://www.fon.hum.uva.nl/praat/>
- [13] FISCHER E.; GOBERMAN A. M. *Voice onset time in Parkinson disease* [online]. Journal of Communication Disorders, č. 43, 2010: s. 21–34. [cit. 26. 11. 2016]. Dostupné z URL: <https://www.ncbi.nlm.nih.gov/pubmed/19717164>
- [14] BLOMGREN M.; GOBERMAN A. M. *Fundamental frequency change during offset and onset of voicing in individuals with Parkinson disease.* [online]. Journal of Communication Disorders, ročník 22, č. 2, 2008: s. 178–191. [cit. 27. 11. 2016]. Dostupné z URL: <https://www.ncbi.nlm.nih.gov/pubmed/16950600>
- [15] GOBERMAN A. M.; ELMER LW. *Acoustic analysis of clear versus conversational speech in individuals with Parkinson disease.* [online]. Journal of Communication Disorders, ročník 38, č. 3, 2005: s. 215-230. [cit. 1. 12. 2016]. Dostupné z URL: <https://www.ncbi.nlm.nih.gov/pubmed/15748725>
- [16] ROUBICKOVA, J.; HEDANEK, J.; STRANIK, A. *Test 3F: Dysartrický profil* [online]. 2012: s. 1–2 [cit. 4. 12. 2016]. Dostupné z URL: <http://www.klinickalogopedie.cz/index.php?pg=aktuality&aid=164>

- [17] MEKYSKA, J. *Analýza řečových promluv pro IT diagnostiku neurologických onemocnění: dizertační práce*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2014.
- [18] GOSLING, J.; JOY, B.; STEELE, G.; BRACHA, G.; BUCKLEY, A. *The Java Language Specification, Java SE 8 Edition*. Addison-Wesley Professional, 2014.
- [19] ROBINSON, M. *Swing*, Manning, Greenwich, 2003
- [20] CONNOLLY, D.; MASINTER, L. *The 'text/html' Media Type* Internet Requests for Comments, RFC Editor, 2000, Dostupné z URL: <https://tools.ietf.org/html/rfc2854>
- [21] LIE, H.; BOS, B.; LILLEY, C. *The text/css Media Type* Internet Requests for Comments, RFC Editor, 1998, Dostupné z URL: <https://tools.ietf.org/html/rfc2318>

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

AR	Articulation Rate - artikulační rychlost
DDK	Diadochokinetic tasks - diadochokinetické úlohy
DSP	číslicové zpracování signálů – Digital Signal Processing
HD	Hypokinetická dysartrie
HTML	HyperText Markup Language
NST	Net Speech Time - doba čistě řečového signálu
PARCZ	Czech Parkinsonian Speech Database
PN	Parkinsonova nemoc
PR	Pause Ratio - procentuální poměr pauz
SD	Standard Deviation - směrodatná odchylka
TKEO	Teager-Kaiser Energy Operator - Teagerův-Kaiserův energetický operátor
TPT	Total Pause Time - celkové trvání pauz
TSR	Total Speech Rate - absolutní tempo řeči
TSR	Total Speech Rate - absolutní tempo řeči
TST	Total Speech Time - absolutní trvání řečového signálu
VAI	Vowel Articulation Index - artikulační index samohlásek

SEZNAM PŘÍLOH

A Obsah přiloženého CD

56

A OBSAH PŘILOŽENÉHO CD

Na přiloženeém CD lze nalézt spustitelný *jar* soubor výsledného programu. Níže lze vidět adresářovou strukturu obsahu CD.

```
/ ..... Kořenový adresář přiloženého CD
├── SpeechPROCSwing/.....Kořenový adresář projektu se zdrojovými soubory
├── SpeechPROC.jar ..... Spustitelný soubor aplikace
└── DP/ ..... Diplomová práce včetně zdrojových souborů LATEX
```